NOAA Technical Memorandum NMFS F/NWC-197

Modeling Dynamic Fish Populations

by
Keith R. Criddle

April 1991

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service

BIBLIOGRAPHIC INFORMATION

PB91-202770

Report Nos: NOAA-TM-NMFS-F/NWC-197

Title: Modeling Dynamic Fish Populations.

Date: Apr 91

Authors: K. R. Criddle.

Performing Organization: National Marine Fisheries Service, Seattle, WA. Alaska Fisheries Science Center.**Alaska Univ., Fairbanks.

Type of Report and Period Covered: Technical memo.

Supplementary Notes: Prepared in cooperation with Alaska Univ., Fairbanks.

NTIS Field/Group Codes: 48B, 98F

Price: PC A05/MF A01

Availability: Available from the National Technical Information Service, Springfield, VA. 22161

Number of Pages: 82p

Keywords: *Fish populations, *Population dynamics, *Fish management, *Fisheries, *Mathematical models, Approximation method, Yield equations, Fish stocking, Forecasting, Time series analysis, Statistical analysis, Least squares method, Fish harvest.

Abstract: Uncertainty about the nature and significance of nonlinearities and the manner in which dynamics affect future realizations makes model specification the most difficult aspect of modeling fish population dynamics. The paper shows that nonlinear dynamic systems can be approximated by dynamic polynomial expansions. The residuals of these dynamic approximations include omitted higher degree terms and higher order dynamics, and as a result, may be nonlinear and will be characterized by complex serial correlations. Because even nonlinear processes have linear state-space representations, a state-space approach can be used for data-based specification of the dynamics of these residuals. A recently introduced multivariate state-space time-series modeling algorithm, system-theoretic time-series, can be used to create a particularly robust state-space representation. Combining the model approximation methodology with time-series modeling of the residuals results in accurate models of short-run dynamics. The approximate structural modeling methodology is illustrated with single species models of yellowfin sole (Limanda aspera) biomass, and multispecies models of Pacific halibut (Hippoglossus stenolepis), yellowfin sole, and red king crab (Paralithodes camtschatica) biomasses in the eastern Bering Sea. A final application of the approximate structural/system-theoretic time-series model includes walleye pollock (Theragra chalcogramma), Pacific halibut, Pacific cod (Gadus macrocephalus), yellowfin sole, and red king crab.

MODELING DYNAMIC FISH POPULATIONS

by

Keith R. Criddle*

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE, BIN C15700
Seattle, WA 98 115-0070

*Department of Economics
University of Alaska Fairbanks

April 1991

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Uncertainty about the nature and significance of nonlinearities and the manner in which dynamics affect future realizations makes model specification the most difficult aspect of modeling fish population dynamics. This paper shows that nonlinear dynamic systems can be approximated by dynamic polynomial expansions. The residuals of these dynamic approximations include omitted higher degree terms and higher order dynamics, and as a result, may be nonlinear and will be characterized by complex serial correlations. Because even nonlinear processes have linear state-space representations, a state-space approach can be used for data-based specification of the dynamics of these residuals. A recently introduced multivariate state-space time-series modeling algorithm, system-theoretic time-series, can be used to create a particularly robust state-space representation. Combining the model approximation methodology with time-series modeling of the residuals results in accurate models of short-run dynamics.

The approximate structural modeling methodology is illustrated with single species models of yellowfin sole (Limanda aspera) biomass, and multispecies models of Pacific halibut (Hippoglossus stenolepis), yellowfin sole, and red king crab (Paralithodes camtschatica) biomasses in the eastern Bering Sea System-theoretic time-series methods are demonstrated in a series of regional models of Pacific halibut biomass. The single-species models of Pacific halibut biomass, and the multispecies models of yellowfin sole, Pacific halibut, and red king crab biomass are later used to illustrate the approximate structural/system-theoretic time-series mode. A final application of the approximate structural/system-theoretic time-series model includes walleye pollock (Theragra chafcogramma), Pacific halibut, Pacific cod (Gadus macrocephalus), yellowfin sole, and red king crab.

THIS PAGE INTENTIONALLY LEFT BLANK

CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

INTRODUCTION

Although management objectives differ widely from one fishery to the next, harvest levels are invariably set in accordance with expectations about future abundance. In order to successfully manage dynamic fish stocks, fisheries management agencies must have accurate forecasts of the time paths of key decision variables such as abundance. Uncertainty about the nature and significance of nonlinearities and the manner in which dynamics affect future realizations makes model specification the most difficult aspect of modeling dynamic fish populations. Although coefficient estimation may be difficult given limited observations on the system, it is model misspecification that biases their estimates.

Many different structural models have been used to predict abundance and to determine optimal harvest levels; however, none has proven overwhelmingly successful. Often these models are mutually exclusive due to various restrictions implied by competing specifications--primarily nonlinearities, interspecies interactions, and higher order dynamics. It is not that the relevant variables are unknown, rather they are unobservable or it is not known how they enter the system. Although biologists agree that the fundamental drivers of population dynamics are natality, growth, fecundity, reproduction, and mortality, the structural models that have been proposed are largely mutually exclusive and often fail to provide the precise short-term forecasts required by fisheries management agencies.

Traditional approaches to modeling dynamic systems suffer from a need to specify both model structure and dynamics a priori. Several difficulties typically arise in the estimation of structural relationships from observed harvest rates and effort. These include biases introduced by the inclusion of unobservable variables, the omission of important exogenous variables, and the inclusion of lagged dependent variables. Uhler (1979) remarks that because models of fish population dynamics include latent variables (e.g., population abundance), parameter estimates will be biased even in large samples. Berck and Johns (1985) correct for latent variables through Kalman filtering. When environmental variation is nonrandom, Armstrong and Shelton (1988)

show that intrinsic growth rates and density dependence are either overstated or understated depending on the relationship between the mean age of the spawning stock and the periodicity of the forcing environmental variation. Because dynamic models of biological systems include lagged dependent variables, Mendelssohn (1980) notes that coefficient estimates will be biased unless the estimation procedure corrects for autocorrelation in the errors. Walters (1985 and 1987) suggests that additional biases result from the correlation between current residuals of the stock-recruitment relationship and the subsequent spawning stock, and from the nonstationarity of the structural components through time. To these can be added the problem of model misspecification, both in relations between variables and in the specification of model dynamics.

The mutual inconsistency of common structural models combined with problems of model specification suggests three alternative approaches: approximate structural modeling, time-series modeling, and approximate structural/time-series modeling. In approximate structural modeling, model specification is approached as the formal dynamic approximation of an encompassing representation of the system. The residuals of these dynamic approximations include omitted higher (degree terms and higher order dynamics, and as a result, although these approximations are stationary, they may be nonlinear and will be characterized by complex serial correlations.

Forecasts of fish abundance have traditionally been based on structural models. However, when the dynamics of a process are important relative to static relations with other variables, it can be more accurate to use the final forms of the equations of interest rather than a complete structural model (Zellner and Palm 1974). In other words, simple time-series models may provide more accurate forecasts than complex ecosystem models. Several recent papers have applied time-series modeling techniques to stock estimation. Saila, Wigbout, and Lermit (1980) consider three time-series models in an application to 12 years of monthly rock lobster ***(Jasus edwardsii)*** landings: monthly averages, harmonic regression analysis, and autoregressive integrated moving average (ARIMA) models. Their results favor an ARIMA model, although their harmonic regression model fits equally well in sample used for coefficient estimation. Mendelssohn (1980) estimates ARIMA and transfer function models for skipjack tuna ***(Euthynnus pelamis)*** landings. Noakes et

al. (1987) contrast structural Ricker stock-recruitment models with ARIMA and transfer function models in the estimation of North Sea herring ***(Clupea harengus)*** and Pacific halibut ***(Hippoglossus stenofepis)*** fisheries. Both time-series models are preferred to the Ricker structural model. The importance of auxiliary time-series is also emphasized by Fogarty (1988), who reports 10-27% reductions in the variance of transfer function models over pure ARIMA models of lobster catch. A multivariate time-series model of six commercial fisheries in Lake Superior is reported in Cohen and Stone (1987). Although time-series methods can provide improved representation of model dynamics, model specification remains problematic for traditional time-series methods.

Because even nonlinear processes have linear state-space representations, a state-space time-series approach can be used for data-based specification of the important nonlinear terms. A recently developed multivariate state-space time-series modeling algorithm is reviewed and extended later in this paper. The results of the state-space representation are considerably improved multispecies biomass forecasts.

Approximations to the encompassing structural model and state-space time-series methods can be integrated to form a model that is able to detemine model dynamics based on the rank of a matrix of autocorrelations. Applying system-theoretic time-series techniques to the residuals of the approximate encompassing model consistently produces reliable forecasts of biomass.

APPROXIMATE STRUCTURAL MODELING

Even if the precise nature of a nonlinear dynamic system is unknown, it may be possible to develop approximations to the true model if the most important variables are known. Careful specification of these approximations can result in simple models which closely emulate the behavior of the system.

Approximate Population Dynamics

The lack of a compelling model of population dynamics might suggest that the processes involved are unknown, whereas in fact, they are well understood. The abundance of a cohort (age

group or year class) within a natural population declines monotonically through time. The sum of abundance across cohorts is total abundance and will vary through time as a function of various mortality rates and the rate of recruitment. Biomass is the weight of a cohort, and changes through time as a result of metabolic growth, predation, disease and other natural mortality, spawning, and harvesting. Except during periods of starvation, biomass strictly increases through time as a function of metabolism, and declines as a function of predation, harvesting, and natural senescent mortality. Spawning exerts a seasonal influence on cohort biomass. The net impact of growth, predation, harvesting, and natural mortality is that cohort biomass first increases, and then decreases through time. Total biomass (like total abundance) is a sum across cohorts which varies periodically over time.

The dynamics of a population are approximately continuous for large numbers of individuals and are often represented as differential equations. Alternatively, because increments and decrements to populations are discrete, their dynamics can be represented as a vector-difference equation,

$$x_t = x_{t-1} - m_{t-1} - p_{t-1} + r_{t-1} - h_{t-1} . \tag{1}$$

Current abundance (x_t) is determined by lagged abundance (x_{t-1}), recruitment (r_{t-1}), predation (p_{t-1}), natural mortality (m_{t-1}), and harvesting (h_{t-1}).

The difficulty in specifying population dynamics is that even if it is assumed that harvesting and biomass are observable, natural mortality, predation, and recruitment are not. Nevertheless, they are related to relative abundance in the present and recent past periods:

$$\begin{aligned} m_{t-1} &= \beta_1 x_{t-1} + \zeta^m_{t-1}, \\ p_{t-1} &= \beta_2 x_{t-1} + \zeta^p_{t-1}, \text{ and} \\ r_{t-1} &= \sum_{i=1}^{\ell} \beta_{2+i} x_{t-i} + \zeta^r_{t-i} . \end{aligned} \tag{2}$$

The first relation expresses natural mortality as a function of own-lagged biomass; that is, B_1 is a diagonal matrix. The coefficient matrix (B_2) specifies predation as a function of the immediately lagged abundance of all species. The coefficient matrices (B_{2+i}) specify recruitment as a function of the biomass of all species from t-l until the present, where l is the maximum lag on the individual series. After substituting the relations in Equation (2) into Equation (l), the model becomes,

$$x_t = [I - \beta_1 - \beta_2]x_{t-1} + \sum_{i=1}^{\ell} \beta_{2+i} x_{t-i} - h_{t-1} + u_t . \tag{3}$$

After simplifying notation, Equation (3) can be written as:

$$x_t = \sum_{i=1}^{\ell} \beta_i x_{t-i} - h_{t-1} + u_t . \tag{4}$$

Equation (4) is a linear matrix difference equation of order l with a potentially complicated error structure.

Generalizing equation (4) to allow for nonlinearities results in the encompassing model,

$$x_t = f(x_{t-1}, x_{t-2} + \ldots + x_{t-\ell}) - h_{t-1} + u_t , \tag{5}$$

where f(x) is the net-growth function and measures the present impact of past recruitment, predation, and competition. Even though x_t may have a complex serial correlation structure, the joint distribution of the data is invariant with respect to t. Equation (5) can be generalized to embed single and multispecies age-structured models using transfer matrices like those used in Leslie-matrix models. In multispecies age-structured models, x_t is an SC element vector of abundances of the c-age groups included in the s-species.

Most commercial fishery harvests are limited by quotas or season length. When harvest limits or bycatch quotas are binding, catch is predetermined. If catch limits are nonbinding, then fishing mortality can be modeled as a function of abundance; catchability; and fishing effort, which in turn is a function of expected profits.

Because the form of the net-growth function and the length of the maximum system lag l are unknown, Equation (5) cannot be estimated directly. In traditional structural models, restictions are arbitrarily imposed on the coefficients in Equation (5) to make estimation feasible. Alternatively, the restrictions implied by the various model specifications can be nested as constraints on terms in formal dynamic approximations to the process.

One such approximation is a dynamic Taylor-series expansion, that is, a polynomial expansion of a process with distributed lags. To simplify notation, define the s(l+1) element column vector X as

$$\mathbf{X} = [\mathbf{x}_t', \mathbf{x}_{t-1}', \ldots, \mathbf{x}_{t-\ell}']',$$

where the prime denotes transposition. A second-degree Taylor-series expansion of the model in Equation (5) is

$$f(\mathbf{x}) = f(\mathbf{X}_0) + \mathcal{J}[\mathbf{X}-\mathbf{X}_0] + \frac{1}{2}\{[\mathbf{X}-\mathbf{X}_0] \otimes \mathbf{I}_s\}' \mathcal{H}[\mathbf{X}-\mathbf{X}_0] + \xi_t, \qquad (6)$$

where the Hessian (H) is an $s^2(l+1) \times s(l+1)$ matrix of the first derivatives of vec(J) with respect to X, and the Jacobian (J) is an s * s(l+1) matrix whose ij-th element is the derivative of the i-th species' dynamics equation with respect to the j-th element of X. [1] The Jacobian is composed of an $s \times s$ identity matrix and l $s \times s$ dynamic matrices

$$\mathcal{J} = [\mathbf{I}_s, \nabla f_1, \nabla f_2, \ldots, \nabla f_\ell],$$

with the derivatives of f with respect to x_{t-i} denoted ∇_{fi} . Cubic and higher degree nonlinearities are summed in the s-element approximation error vector (ξ_t) which is of order $|X-X_0|^{r+1}$ where r is the power of the highest degree term included in the approximation. In general ξ_t , is a serially correlated nonlinear function of the stationary stochastic process, X, and is itself stochastic.

[1]The operator vec() stacks successive columns of a matrix below one another; $\otimes$ is the Kronecker product.

Alternative structural models can be represented as sets of restrictions on the terms in these expansions, with the most important differences appearing in the form of the second term in the first-degree approximate model

$$f(X) = f(X_0) + \mathcal{J}[X-X_0] + \xi_t, \tag{7}$$

with an approximation error ξ_t of order $|X-X_0|^2$. A more explicit representation of Equation (7) is

$$x_t = \beta_0 + \nabla f_1 x_{t-1} + \nabla f_2 x_{t-2} + \ldots + \nabla f_\ell x_{t-\ell} - h_{t-1} + \xi_t. \tag{8}$$

Equation (8) can be viewed as a statistical model or as an encompassing structural model. As a statistical model, the coefficients matrices are estimated and tested to reveal the significance of various interactions. Because the off-diagonal terms in the ∇f_i represent interactions between species, where the off-diagonal terms are insignificant for all i, interspecific dynamics are unimportant and the stocks can be managed independently. Trophic hierarchies are indicated by a block-recursive structure in one or more ∇f_i.

Structural models can be motivated as restrictions on the terms in one or more of the derivatives off: For the discrete-time logistic surplus production model,, ∇f_1 is a diagonal matrix with coefficients linear in x_{t-1}, and $\nabla f_i = 0$ for all $i \neq 1$. Delay-difference models with discrete logistic stock-recruit and growth functions resemble logistic surplus production models, but also include a second diagonal matrix ∇f_ℓ composed of coefficients linear in x_{t-1}, and $\nabla f_i = 0$ for all $i \neq 1$ or l. In multispecies models, the ∇f_i may be nondiagonal for at least some i to reflect predation and competition. The Appendix includes derivations of some single and multispecies models nested by Equation (5).

While common structural models disagree over the specification of nonlinear components in f and the length and relevance of higher order dynamics, each includes linear components and first order dynamics. This suggests that Equation (8) could be rewritten as a linear first order difference equation in x_t and a formal approximation error ε_t with autoregressive and moving average components:

$$x_t = \beta_0 + \beta_1 x_{t-1} - h_{t-1} + \varepsilon_t \text{ and} \quad (9)$$

$$\varepsilon_t = [\nabla f_1 - \beta_1] x_{t-1} + \nabla f_2 x_{t-2} + \ldots + \nabla f_\ell x_{t-\ell} + \xi_t, \quad (10)$$

where ξ_t is the stochastic approximation error introduced above. Because ξ_t, is a stationary stochastic process, ε_t is simply a zero-mean stochastic error with a complicated serial correlation structure. If nonlinearities and higher order dynamics were known, exact structural models could be specified. Otherwise, formal approximations of the encompassing model which do not impose arbitrary restrictions may provide superior forecasts.

Statistical Tests of the Degree of Nonlinearity

Classical tests to determine the degree of polynomial distributed-lags estimate higher degree models first, and then later evaluate the significance of the coefficients of the higher degree terms. If the coefficients on the higher degree nonlinear terms are insignificant, a lower degree model is estimated. The limitations of this approach are that there is no assurance that higher degree terms excluded from the most general model estimated are, in fact, insignificant, and degrees of freedom are lost in estimating coefficients on included higher degree nonlinear terms which may not be significant and are potentially collinear. The expansions of Equation (5) suggest a pretest of model degree which estimates the significance of excluded higher degree nonlinear terms. Note that the true population dynamics model, Equation (5), does not include an intercept. That is, there is no constant level about which abundance varies; rather it is everywhere determined by additions, removals, and immediate past levels. However, Taylor-series, Runge-Kutta, and other approximations of n-th degree systems include intercepts in all expansions through (n-1)-th degree, that is, higher degree nonlinearities appear in an intercept when they are excluded from the Taylor-series expansions. For illustration, consider the discrete logistic model

$$x_t = [1+g]x_{t-1} - \frac{g}{k}x_{t-1}^2 - h_{t-1},$$

a quadratic equation in x_{t-1} without an intercept. The first-order Taylor-series linear approximation of this model excludes 'quadratic terms in x_{t-i} and includes an intercept

$$x_t \equiv \frac{g}{k}x_0^2 + \{[1+g] - \frac{2x_0}{k}]\}x_{t-1} - h_{t-1} .$$

The approximation error ξ_t, can be determined by subtracting the linear expansions from the true model

$$\xi_t = \{x_t - [1+g]x_{t-1} + \frac{g}{k}x_{t-1}^2 - h_{t-1}\} - \{x_t - [1+g]x_{t-1} + \frac{2g}{k}x_0x_{t-1} - x_0^2 - h_{t-1}\}$$

$$= \frac{g}{k}[x_{t-1}^2 - 2x_0x_{t-1} + x_0^2] .$$

The omitted higher degree terms are approximated by introducing an upward bias equal to $2x_0\frac{g}{k}$ in the estimate of B_1, and including an intercept equal to $\frac{g}{k} + x_0^2$. Because x_0^2 is greater than $2x_0$, the share of ξ_t, included in the intercept will exceed the portion included in B_1.

Although the degree of the encompassing model is unknown, the highest degree supported by the data can be determined from a t-test on the significance of the intercept of the approximate model.

Because the omitted higher order dynamics of the system will be estimated from the residuals of the approximate first-order model using a state-space time-series technique presented later, there is no loss in generality to specifying first order dynamics in the approximate model. The unique structure of equations used to represent discrete population dynamics permits a one-dimensional search for model degree; this is in contrast to the usual case which requires a two-dimensional search (see e.g., Pagano and Hartley 1981). As with any sequential test, the power of the test declines as the number of iterations increases. The significance of the i-th test in a sequence of tests is

$$1 - \prod_{j=1}^{i}(1 - \alpha_j),$$

where α_j is the level of significance of the j-th individual test (Trivedi and Pagan 1979).

Yellowfin sole

Table 1 includes coefficient estimates and corresponding t-ratios for first-, second-, and third-degree Taylor-series expansions of Equation (5) for a single-species model of yellowfin sole ***(Limanda aspera)*** biomass in the eastern Bering Sea. The data are cohort analysis estimates of the abundance of age 7+ yellowfin sole for 1959 through 1986 (Bakkala and Wespestad 1986, 1988).

The first-degree approximate model includes an intercept which has more than a 99.95% probability of being different from zero, suggesting that the true model is characterized by at least second-degree nonlinearities. Coefficient estimates on x_{t-1} and x_{t-1}^2 in the quadratic model are significant at the 99.9 and 99.5% levels, respectively. Because the intercept of the quadratic model is not significantly different from zero, the second-degree approximation cannot be distinguished from the true model on the basis of the data. In fact, it is found that only the linear term in the cubic model has a significant probability of not occurring by chance alone.

Analysis of the dynamic properties of the three estimated models indicates that the linear model is explosive, while the quadratic and cubic models converge monotonically towards equilibria of 2,283,000 and 2,173,000 metric tons (t), respectively.

Table 2 compares the performance of the linear, quadratic, and cubic models of yellowfin sole biomass. Although performance on most criteria improves as model complexity increases, the absolute differences are small, and information content of the cubic model as measured by Schwartz's criterion is less than that of the quadratic model. Figure 1 presents forecasts from the linear, quadratic, and cubic models along with the actual observations and the residuals. The three graphs are similar, suggesting that there is little gain from estimating cubic or higher degree models (as indicated by the intercept test), with possibly inferior out-of-sample forecasts due to increases in sampling error. Model performance is primarily determined by the linear terms, with higher degree terms modestly reducing serial correlation of the errors.

Table l.-- Approximate models of yellowfin sole biomass: Coefficients.[a]

Model	Variables			
	Intercept	x_{t-1}	x^2_{t-1}	x^3_{t-1}
Linear	154.58[b]	1.020[b]		
	(4.22)	(30.96)		
Quadratic	-28.03	1.472[d]	-2.0×10^{-4}[c]	
	(-0.41)	(9.77)	(-3.05)	
Cubic	65.16	1.127[e]	1.4×10^{-4}	-9.8×10^{-8}
	(0.42)	(2.08)	(0.27)	(-0.66)

[a]Biomass of yellowfin sole is in 1000 t. T-ratios are reported in parentheses below each coefficient.
[b]Significant at $\alpha = 0.0005$.
[c]Significant at $\alpha = 1 - (1-.0005)(1-.0005) = 0.001$.
[d]Significant at $\alpha = 1 - (1-.0005)(1- 0.005) = 0.005$.
[e]Significant at $\alpha = 1 - (1-.0005)(1- 0.005)(1-.025) = 0.030$.

Table 2. -- Approximate models of yellowfin sole biomass: Statistics.

	RMSE	%RMSE	MAD	R^2	SC	Q
Linear	94.71	9.99	75.74	.973	9.352	36.06
Quadratic	80.91	8.54	66.71	.980	9.163	7.63
Cubic	80.13	8.45	64.22	.981	9.269	7.61

RMSE is root-mean-squared error; %RMSE is percentage root-mean-squared error; MAD is mean absolute deviation: R^2 is the coefficient of determination; SC is Schwartz's criterion; and Q is the the Ljung-Box portmanteau test.

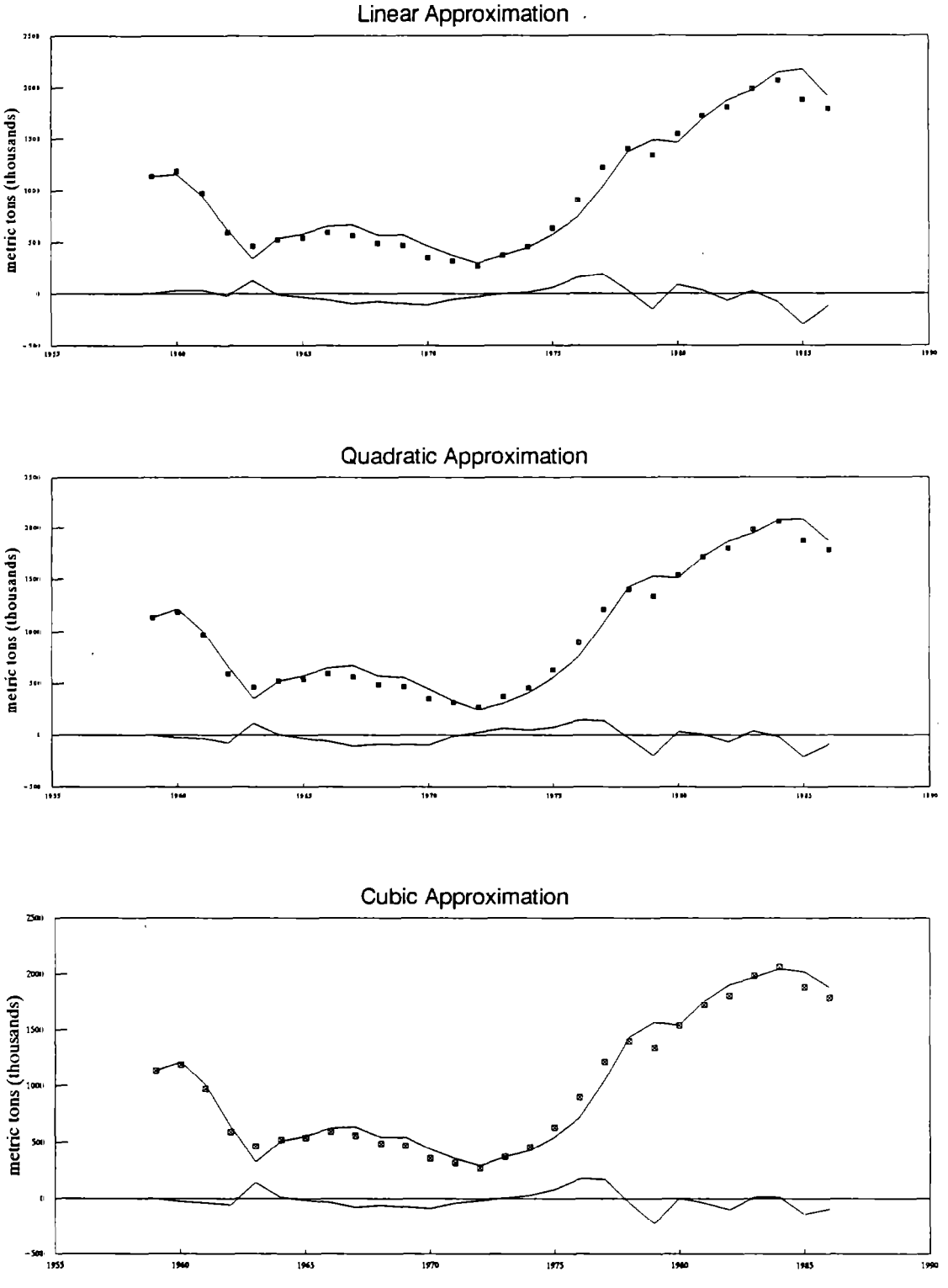

Figure 1 .--First-, second-, and third-degree models of yellowfin sole biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

Constrained Indirect Least Squares (CIDLS)

One of the limitations faced by researchers is the short time-series of data available for stock-assessment studies. In order to conserve degrees of freedom, it is desirable to reduce the number of coefficients to estimate. However, because of the complexity of ecological linkages, it is uncertain which interactions are inconsequential. Equation (9) approximates the encompassing model with a matrix first order difference equation. The approach discussed here may further restrict the number of coefficients required. The coefficient matrix Bi represents first-order interactions among the species modeled. If unimportant interactions are restricted to zero, then the remaining coefficients can be estimated with increased degrees of freedom.

Wegge (1978) estimates overidentified models by sequentially pretesting the overidentifying restrictions using constrained indirect least squares (CIDLS). The CIDLS estimates of the exactly identified model that have the same asymptotic properties as full-information maximum-likelihood (MLE) estimates. Because model determination is conditional on the choice of identifying restrictions, selection of the identifying restrictions is critical.

Discrete-time models of population dynamics offer a natural set of exactly identifying restrictions. Because interactions affect the model only after a lag of at least one period, there is an identity matrix representing contemporaneous interactions which premultiplies the left hand side of Equation (9) that constitutes the set of exactly identifying restrictions.

Sequential pretesting of overidentifying restrictions constrains the likelihood function. At each iteration, the restriction which least reduces the likelihood function is imposed so long as the Wald statistic with the addition of the restriction does not exceed the critical Chi-squared value.

Pacific Halibut, Yellowfin Sole, and Red King Crab

The incidental harvest of ***red king*** crab ***{Paralithodes camtschatica)*** and Pacific halibut by trawl fisheries for yellowfin sole is of concern to the North Pacific Fisheries Management Council (NPFMC). Recent increases in the harvest of yellowfin sole and other trawled species, accompanied by decreases in the abundance of red king crab, has led crab fishermen to lobby the

Council for tighter restrictions on incidental catches of crab by trawlers. Trawlers' associations counter that percentage bycatch is already low, and that further reductions would substantially reduce the profitability of trawling. Accurate models of population dynamics are integral to the resolution of these conflicting claims.

Pacific halibut are reported to prey on adult red king crab, while yellowfin sole prey on the juveniles. It is most likely that increases in the abundance of sole will result in decreases in crab abundance, and that increases in crab abundance will be followed by increases in sole abundance. However, it is also entirely possible that abundance in each species could be exogenously driven. Because crab and halibut are harvested incidental to the yellowfin sole fishery, increases in abundance accompanied by increases in the catch limits for yellowfin sole might result in increased bycatch of crab and halibut. A literature review of predator-prey linkages can be found in Livingston and Goiney (1983). Predator-prey linkages between the species groups included in the model and those excluded are relatively insignificant[2]. Without clear guidance from biological theory, the CIDLS algorithm can be used to determine appropriate restrictions directly from the data. Recent catch and abundance information is included in Table 3.

Model Estimation--A maximum-likelihood first-order linear model for Pacific halibut, yellowfin sole, and red king crab biomass was estimated using CIDLS (Wegge 1978). The estimated model is

$$\begin{bmatrix} x_t^h \\ x_t^s \\ x_t^c \end{bmatrix} = \begin{bmatrix} 1.01 & 0.81 & -0.01 \\ 0.02 & 1.03 & 0.001 \\ 3.41 & -8.95 & 1.28 \end{bmatrix} \begin{bmatrix} x_{t-1}^h \\ x_{t-1}^s \\ x_{t-1}^c \end{bmatrix} - \begin{bmatrix} h_{t-1}^h \\ h_{t-1}^s \\ h_{t-1}^c \end{bmatrix}.$$

All coefficients are asymptotically significant at $\alpha = 0.05$. The coefficients of determination (R^2) are 0.922, 0.979, and 0.854 for the halibut, yellowfin sole, and red king crab equations, respectively.

[2]Pacific cod (Gadus macrocephalus) are thought to be important predators of red crab, but were not included due to the limited number of observations on their abundance. The omission of cod may introduce bias into the estimates of the parameters of the three equations.

The eigenvalues of the coefficient matrix are 1.080 ± 0.971i and 1.158, indicating that the system is unstable. However, the system could be stabilized by careful selection of the harvest levels. Graphs of actual biomass and estimates based on the MLE coefficients are included as Figure 2.

The CIDLS model permitted the imposition of three overidentifying restrictions. Coefficient estimates for the restricted (CIDLS) model are

$$\begin{bmatrix} x_t^h \\ x_t^s \\ x_t^c \end{bmatrix} = \begin{bmatrix} 1.04 & 0.75 & -0.01 \\ & 1.07 & 0.002 \\ & & 1.32 \end{bmatrix} \begin{bmatrix} x_{t-1}^h \\ x_{t-1}^s \\ x_{t-1}^c \end{bmatrix} - \begin{bmatrix} h_{t-1}^h \\ h_{t-1}^s \\ h_{t-1}^c \end{bmatrix}.$$

All of the coefficients of the CIDLS first order approximate model are statistically significant at the 99.5% level. Overall fit of the model declined by less than 1% in the halibut equation, and by little over 1 and 5% for the sole and crab equations, respectively, even though the number of coefficients was reduced by one-third. The coefficients of determination for this model are 0.917, 0.973, and 0.800 for the Pacific halibut, yellowfin sole, and red king crab equations, respectively. Figure 3 presents CIDLS estimates of halibut, sole, and crab biomass. Red king crab dynamics are a first order linear autoregression. Yellowfin sole dynamics are nearly a linear first order autoregression because the interaction coefficient with crab is very small. Error plots included in Figure 3 indicate the presence of serial correlation as predicted by Equation (10).

Structural stability is determined by the eigenvalues of the dynamic coefficient matrix, which in the case of triangular matrices, are the diagonal elements. Because all three eigenvalues lie outside the unit circle, the solution represents an unstable node.

The statistical properties of the MLE and CIDLS models are compared in Table 4. Halibut biomass is in millions of pounds, red king crab and yellowfin sole are measured in thousand metric tons and million metric tons, respectively. Although the statistics reported in Table 4 indicate that there is little difference between the performance of the two models, the CIDLS model specification is preferred on the basis of Schwartz's Criterion (SC): $SC_{MLE} = 155.16$; $SC_{CIDLS} = 151.04$.

Table 3. -- Biomass and landings of Pacific halibut, yellowfin sole, and red king crab (1000 t).

	Biomass			Landings		
Year	Pacific halibut	Yellowfin sole	Red king crab	Pacific halibut	Yellowfin sole	Red king crab
1969	6.38	469.00	26.28	0.60	167.13	8.65
1970	6.04	353.00	12.29	0.56	133.08	9.19
1971	5.67	317.00	12.77	0.40	160.40	8.16
1972	5.32	273.00	13.26	0.35	47.86	11.89
1973	5.08	371.00	27.49	0.15	78.24	12.79
1974	4.99	456.00	52.25	0.25	42.24	19.62
1975	5.17	631.00	54.41	0.24	64.69	22.59
1976	4.43	895.00	89.18	0.29	56.22	28.66
1977	4.04	1216.00	100.84	0.49	58.37	31.80
1978	3.74	1401.00	122.86	0.61	138.43	39.83
1979	3.34	1339.00	127.71	0.62	99.02	49.01
1980	2.92	1542.00	101.74	0.32	87.39	59.07
1981	2.72	1723.00	32.36	0.54	97.30	15.27
1982	2.96	1809.00	11.96	0.19	95.71	1.36
1983	3.44	1985.00	3.68	2.01	108.39	0.00
1984	4.22	2062.00	7.33	1.44	159.53	1.90
1985	5.31	1880.00	5.91	1.95	227.11	1.90
1986	6.16	1787.70	14.48	2.54	208.60	5.18

Adapted from: Bakkala and Wespestad (1986); Stevens and Macintosh (1986); and Stevens et al. (1988), assuming an average weight of 7 lbs/crab; Quinn et al. (1985); and the International Pacific Halibut Commission (1988).

Table 4. -- Maximum likelihood (MLE) and constrained indirect least-squares (CIDLS) model statistics.

	Mean	RMSE	μ_ξ	%RMSE	MAD
Pacific halibut					
MLE	4.55	0.321	-0.018	7.06	0.240
CIDLS	4.55	0.331	-0.051	7.27	0.270
Yellowfin sole					
MLE	1.14	0.094	0.006	8.23	0.069
CIDLS	1.14	0.106	0.024	9.32	0.080
Red king crab					
MLE	45.38	16.131	-0.130	35.55	12.810
CIDLS	45.38	18.850	2.988	41.54	14.354

RMSE is root-mean-squared error; μ_ξ is the average error; %RMSE is percentage root-mean-squared error, MAD is mean absolute deviation.

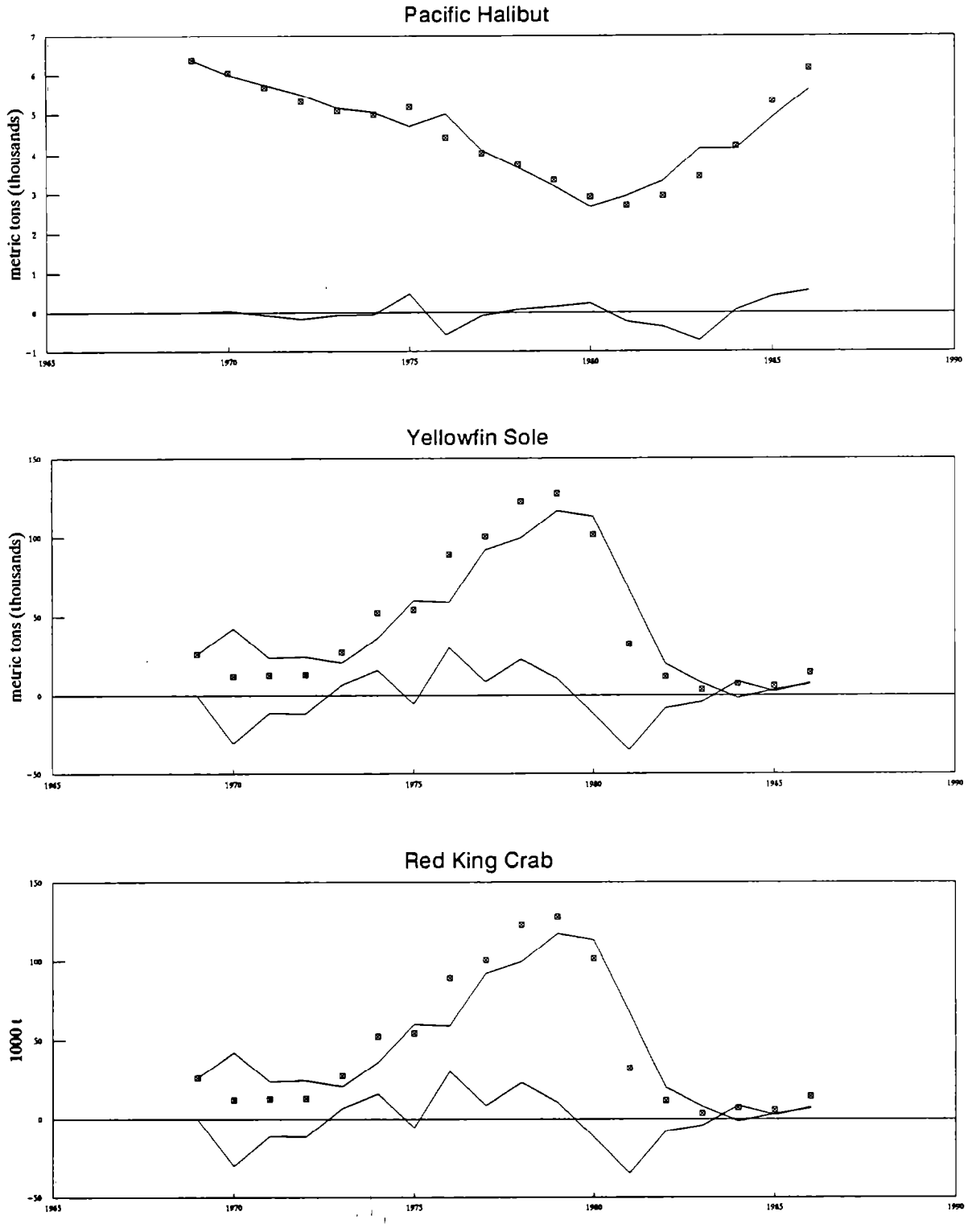

Figure 2. -- Maximum-likelihood model forecasts of Pacific halibut, yellowfin sole, and red king crab biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

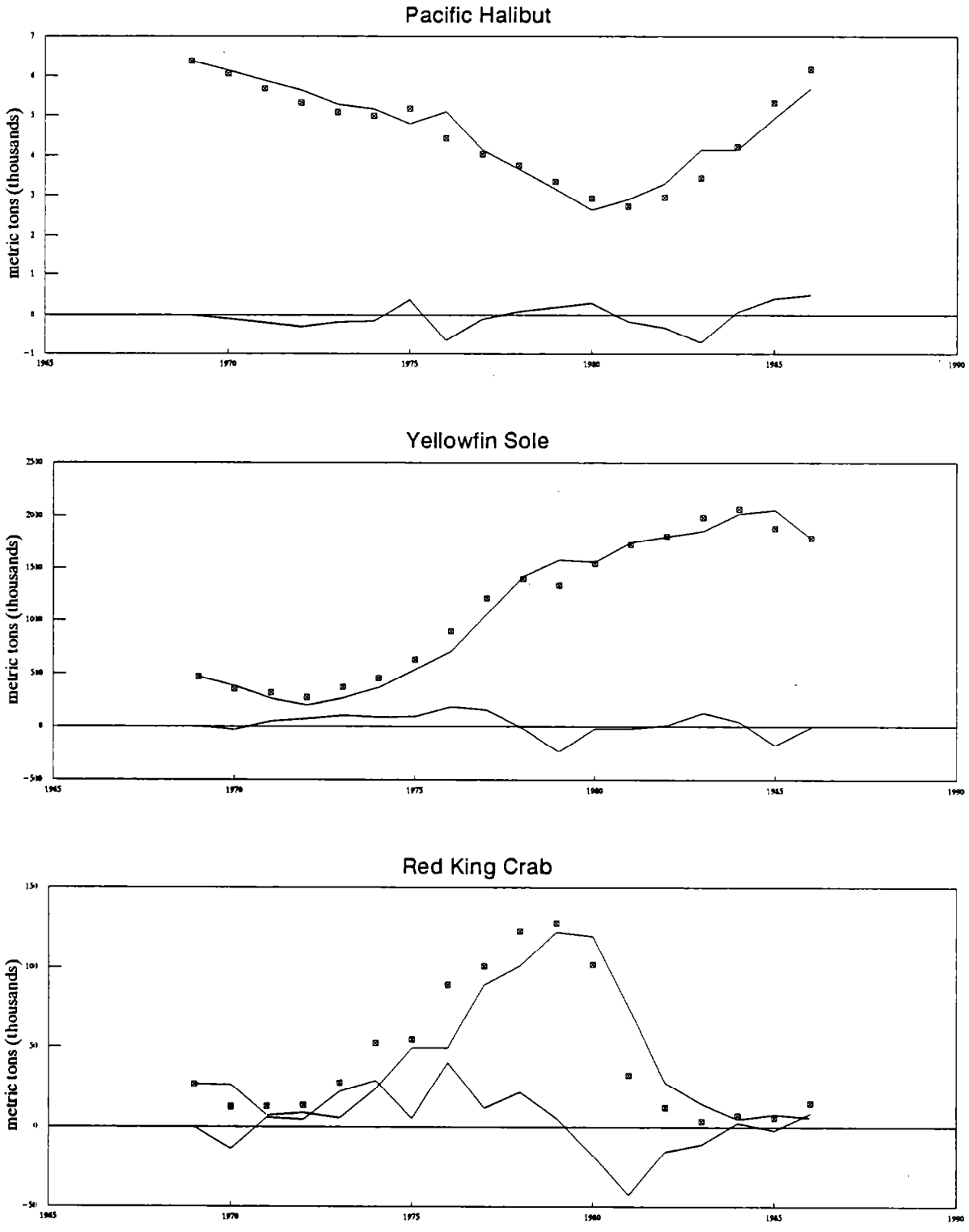

Figure 3. -- Constrained indirect least-squares model forecasts of Pacific halibut, yellowfin sole, and red king crab biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

TIME-SERIES MODELING

All time-series modeling techniques rely on Weld's decomposition which establishes that any stationary stochastic process can be separated into a set of autoregressive terms and another set of serially uncorrelated errors. Many time-series procedures examine the sample autocorrelations and use familiarity with the types of models that might produce these values to specify a model in terms of lagged data values and a moving average of innovations. These regressors are highly correlated with each other almost by definition, leading to large changes in the values of parameter estimates for minor specification changes and making it difficult to determine which of a variety of models will produce the best forecasts. When seasonality is added, the interaction of the seasonal and nonseasonal models adds another level of complexity to the process. The resulting modeling procedure has been criticized both for being highly subjective and for producing finished models that depend on the order of trial of intermediate models.

A new and very powerful multivariate time-series modeling technique based on the principles of linear system theory was introduced by Aoki (1987). This procedure (as well as several others) depends on states, or dynamic factors, to summarize the information in the series rather than using past values and errors. While the state space representation and the classical time-series representations can always be derived from each other, they have different restrictions and generally result in different models when applied to the data, States are minimum sufficient statistics for the past history of the series, and in Aoki's procedure are chosen so that the model specification search is robust with respect to misspecification and is conditionally one-dimensional. Although the development of Aoki's algorithm requires a considerable number of matrix equations, the procedure is easy to apply and produces accurate forecasts.

System-Theoretic Time Series

System-theoretic time series (SITS) uses a unique arrangement of sample autocovariances to determine model dynamics (lag structure) on the basis of matrix rank through appraisal of sample autocorrelation functions. Robustness, with regard to model specification, is obtained by

forming state variables as specific linear combinations of the observed series, thereby permitting a nonsubjective search over possible models, and assuring consistency of included coefficients even if model order is missspecified (Havenner and Aoki 1988c). As originally proposed, the procedure requires stationary processes and is not appropriate for series that include deterministic or stochastic trends. Aoki (1987) and Havenner and Aoki (1988b) adapt and extend Aoki's original model to nonstationary series.

The procedure requires that the data be centered, that is, the means of each series are subtracted. The model presented below deviates from Aoki's prescription by working with autocorrelations rather than autocovariances (see Havenner and Criddle 1989 for a development of the effects on the procedure of switching to autocorrelations). The model is formed as a pair of matrix equations:

$$Dx_t = D\mu + DCz_{t|t-1} + De_t \quad \text{and} \tag{11}$$

$$z_{t+1|t} = Az_{t|t-1} + [BD^{-1}]De_t\,. \tag{12}$$

There are T observations on each of s stochastic series in x. Equation (11) relates the s-element vector of scaled observations Dx_t to the n-element vector of unobservable states, $z_{t[t-]}$, through the $s \times n$ matrix C. The number of states depends on the degree to which the series follow common processes and on the complexity of the underlying dynamics, with higher order dynamics resulting in increased states. Motion of the state variables is described by Equation (12). Because the number of states is to be determined, there is no loss of generality in assuming that Equation (12) is first-order, as it is always possible to define new elements equal to the lag of other elements in the state vector to reach any lag desired. The coefficient matrix A is an $n \times n$ matrix that relates the state variables to each other through time, while B is an $n \times s$ matrix of Kalman filter updates which incorporates the errors of past-period forecasts on the s-series into future forecasts of the n-states. Given stationarity, the model can be formulated so that the s-element error vector (e_t) is serially uncorrelated and centered on zero with a constant covariance matrix denoted by Ψ. Because the optimal states can be shown to be a linear combination of the observed residuals, Equations (11)

and (12) are written with the same error. The scaling matrix D is a diagonal matrix of inverse standard deviations of the s-series. When the data are centered and scaled by their standard deviations, the autocovariances of the transformed data are also the autocorrelations of the original data. The sample estimate of the lag i autocorrelation matrix is

$$\begin{aligned}\Gamma_i &= E[Dx_{t+i}x_t'D'] \\ &= E[Dx_tx_{t-i}'D'] \\ &= \frac{1}{T}\sum_{t=1}^{T-i} Dx_{t+i}x_t'D' \quad \text{with } i = 0, 1, 2, \ldots 2N .\end{aligned} \tag{13}$$

We want to find a model that adequately approximates this sequence for i = 1, 2, 3, . . .

Because it is developed from a unique arrangement of autocovariances, the system theory approach is able to specify the lag structure of the model on the basis of the rank of an estimated Hankel correlation matrix obtained as the outer product of the scaled, centered observation series. (Hankel matrices are block band counter-diagonal.)

$$H = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \Gamma_3 & \ldots & \Gamma_N \\ \Gamma_2 & \Gamma_3 & & & . \\ \Gamma_3 & & & & . \\ . & & & & . \\ \Gamma_N & & \ldots & & \Gamma_{2N-1} \end{bmatrix} . \tag{14}$$

The r_i are s x s autocorrelation matrices of observations separated by lag i, and 2N-1 is the maximally distant autocorrelation modeled. In practice, selection of N involves trading off errors in the estimated sample autocorrelations at long lags against errors due to truncating the true lag structure[3]. It is important to recall that the determinantal polynomial in the lag operator for any

[3]All modeling procedures require specification of such a parameter, whether explicitly as in this case, or implicitly, as when inspecting the autocorrelations in a more traditional time-series setting.

individual series is the product of the maximum system lag times the number of series, so N may be relatively small for multivariate models.

The Kronecker Theorem demonstrates that the rank of the population counterpart of this matrix is the number of states necessary to characterize the system, n. That is, the dynamic specification of the multivariate time-series model may be determined from the rank of an autocorrelation matrix (Kailath 1980). A computationally accurate method of determining the rank of the Hankel matrix is provided by the singular-value decomposition,

$$H_{n^*} = U_{n^*} \Sigma_{n^*} V'_{n^*} , \tag{15}$$

where Σ_{n^*} is a diagonal matrix of singular values pre- and postmultiplied by the orthonormal singular vectors U_{n^*} and V_{n^*}. The subscript n* indicates that the matrices are based on the n* largest singular values of the Hankel matrix. The essence of the procedure is to approximate the space spanned by the Hankel matrix of autocorrelations by the subspace spanned by those singular vectors associated with nonzero singular values. If the autocorrelations where observed exactly, any singular values beyond the model order would be exactly zero. Because the autocorrelations are sample estimates they contain sampling error, as do the singular values of the Hankel matrix they make up. Ordering the estimated singular values from largest to smallest, any number of rules can be used to consistently estimate the cutoff point (Aoki and Havenner 1989). A criterion analogous to a condition number is the selection of n *, the estimated rank of H, and the required number of states so that

$$\frac{\sigma_{n^*+1}}{\sigma_1} \cong \frac{1}{\sqrt{T}} ,$$

where σ_i is the i-th singular value. Arranging the singular values in order of descending magnitude causes each model to strictly nest all lower order models, thereby assuring consistent estimates of the rank of H, and of common coefficients of z, A, and C even if model dynamics are misspecified (Havenner and Aoki 1988c). In other words, it is hard to get the model wrong.

Equations (11) and (12) can be solved directly for estimates of the matrix of dynamic coefficients A, the coefficient matrix C, and Ω, the matrix of covariances between states and observations:

$$\Omega = E\,[z_{t+1|t}\, x_t' D'] \,. \tag{16}$$

From equations (11), (12), and (16) it can be shown that

$$\Gamma_i = D\, C\, A^{i-1}\, \Omega\, D \,. \tag{17}$$

Applying this definition of Γ_i, the Hankel matrix can be factored into observability O and controllability K matrices:

$$H = \begin{bmatrix} DC \\ DCA \\ DCA^2 \\ . \\ . \\ . \\ DCA^{n-1} \end{bmatrix} \begin{bmatrix} \Omega D & A\Omega D & A^2\Omega D & \dots & A^{n-1}\Omega D \end{bmatrix} = OK. \tag{18}$$

Equating the observability-controllability factorization with the singular-value decomposition, we have two relations based on the n* singular values and associated singular vectors:

$$O_{n^*} = U_{n^*}\, \Sigma_{n^*}^{1/2} \text{ and} \tag{19}$$

$$K_{n^*} = \Sigma_{n^*}^{1/2} V_{n^*}' \tag{20}$$

which can be used to estimate the coefficient matrices A, B, and C.[4] The first block row of H, denoted $H_{,,}$, is used to estimate C. From Equation (18),

[4]The statistical properties of these estimators have been developed in Havenner and Aoki (1988c), where they are interpreted as instrumental variable estimators.

$$H_{1*} = DCK\ , \tag{21}$$

so that an estimate of C is

$$\begin{aligned} C &= D^{-1}\, H_{1*}\, K^{-}_{n*} \\ &= D^{-1}\, H_{1*}\, V_{n*}\, \Sigma^{-1/2}_{n*}, \end{aligned} \tag{22}$$

where D is a sample estimate and K^{-}_{n*} is a singular value generalized inverse of K_{n*}.

'The first block column of H, denoted H_{*1} is used to estimate Ω. Again, from Equation (18), the first block column of the Hankel matrix is

$$H_{*1} = O\, \Omega\, D, \tag{23}$$

so that an estimate of Ω is

$$\begin{aligned} \Omega &= O^{-}_{n*}\, H_{*1}\, D^{-1} \\ &= \Sigma^{-1/2}_{n*}\, U'_{n*}\, H_{*1}\, D^{-1}, \end{aligned} \tag{24}$$

where O^{-}_{n*} is a singular value generalized inverse of O_{n*}.

Estimation of A uses both the relations developed above. Define

$$\overleftarrow{H} = O\, A\, K,$$

where the left arrow notation indicates that $\overleftarrow{H}$ is equivalent to shifting H one block column to the left, deleting the first block column (H_{*1}), and filling in an additional block column on the right (H_{*N+1}). An estimate of A is

$$\begin{aligned} A &= O^{-}_{n*}\, \overleftarrow{H}\, K^{-}_{n*} \\ &= \Sigma^{-1/2}_{n*}\, U'_{n*}\, \overleftarrow{H} V_{n*}\, \Sigma^{-1/2}_{n*}\ , \end{aligned} \tag{25}$$

using both relation (21) and (24).

An estimate of B is obtained from the joint solution of three matrix equations that minimizes Ξ.The matrix equations are derived from the contemporaneous covariance of the state variables,

$$\Xi = E[z_{t+1|t} z'_{t+1|t}]; \tag{26}$$

the covariance of the observation errors,

$$\Psi = E[De_t e'_t D']; \tag{27}$$

the covariance of the states with the observations,

$$\Omega = E[z_{t+1|t} x'_t D']; \tag{16}$$

along with the unconditional data covariance matrix,

$$\Gamma_0 = E[Dx_t x'_t D'] . \tag{13}$$

Using the orthogonality of the innovations e_t and the current states $z_{t|t-i}$, and the fact that both the errors and the states have expected values of zero, taking the expectation of Equation (12) with itself transposed gives

$$\Xi = A\Xi A' + [BD^{-1}]\Psi[BD^{-1}]' . \tag{28}$$

Similarly, taking the expectation of Equation (11) multiplied by its transpose gives

$$\Gamma_0 = D\mu\mu' D' + DC\Xi C' D' + \Psi . \tag{29}$$

The expectation of Equation (12) times the transpose of Equation (11) is

$$\Omega = A\Xi C' D' + [BD^{-1}]\Psi . \tag{30}$$

These equations can be solved for estimates of Ξ, Ψ, and B conditional on the estimates of A, Γ_0, Ω,C, and D. Iterative and noniterative methods for solving Equations (28), (29), and (30) are described in Aoki (1987).

With estimates of A, B, C, and an estimate of the initial state vector z_0, Equations (11) and (12) can be solved for in- and out-of-sample forecasts. The initial conditions can be set to zero since their effect declines asymptotically for stationary processes. Alternatively, the initial states can be backcast or smoothed; see Aoki and Havenner (1989) for details.

Pacific Halibut

Pacific halibut is an important fishery, with 1986 landings valued at $100 million ($84 million in the United States and $16 in million Canada). The fishery is managed for the United States and Canada by the International Pacific Halibut Commission (IPHC). The United States is the sole harvester in the Gulf of Alaska and eastern Bering Sea (IPHC Regulatory Areas 3 and 4), and shares the allowable harvest in Area 2 on a 60:40 basis with Canada. The IPHC is mandated to maintain Pacific halibut stocks at or near the maximum sustainable physical yield.[5] In this capacity, the IPHC establishes gear restrictions, minimum catch size (i.e., size of retainable fish), and sets an annual quota.

Whether stock is viewed in terms of abundance (number of fish) or biomass (weight of fish), it appears to have long natural cycles, perhaps because halibut are long-lived (to 20+ years, with some growing very old) and slow to mature (approximately 8 years for the male and 12 years for the female).

Models and Forecasts-The interactions between Pacific halibut in separate regulatory areas determine whether the best forecasting models are specific to each region or joint over all regions. In unreported investigations, univariate models were preferred to multiregional models. Even contiguous areas and subareas add more sampling error from joint modeling than is removed by exploiting interactions across regions. This section presents several univariate models at increasing levels of disaggregation beginning with a model of coastwide halibut biomass and proceeding through separate models for each regulatory area to the most disaggregated models at the regulatory subarea level. A basic premise of time-series analysis is that the series of interest should be

[5]See Bell (1981) for a comprehensive history of the IPHC and the Pacific halibut fishery.

modeled directly, rather than constructed from components, in order to exploit cancellations and common factors in the data. Besides providing models useful in their own right, the aggregated models provide checks on the forecasts of the less aggregated component series.

The models were specified and estimated using 53 annual observations on the biomass of Pacific halibut for IPHC Regulatory Areas 2, 3, and 4, and for subareas of 2 and 3. Catch limits are currently set for five subregions in Area 4; however, reported biomass estimates are for Area 4 as a whole.

Observations for 1935-84 are reported in Quinn et al. (1985); revisions and additional observations for 1975-87 were obtained from the IPHC. Although the boundaries of the regulatory areas have changed through time, the data are adjusted to 1985 definitions. All model specifications and parameter estimates are based on 45 observations from 1935 through 1979, reserving the 8 observations from 1980 to 1987 for out-of-sample validation. In all cases, the lag parameter N was set to two so that autocorrelations to lag three are included in the Hankel matrix used to determine model order and autocorrelations to lag four are used in estimation.

Coastwide Models and Forecasts--Aggregating biomass over Regulatory Areas 2, 3, and 4 creates a coastwide biomass series with a mean of 259.5 million pounds and a standard deviation of 9.3 million pounds in-sample. Using this mean as a (naive) forecast results in a root-mean-squared error of 58.6 million pounds for the 8 periods out-of-sample, while the model root-mean-squared error is 6.7 million pounds. The coastwide biomass model results are reported in Table 5. The sample mean and variance used in this transformation have been supplied in each case to permit linking the coefficients to the raw data. In- and out-of-sample goodness of fit statistics on the original data (after reversing the centering and scaling transformations) are also presented. Figure 4 plots the model forecasts and the actual data. The vertical bar at 1979 marks the end of the data used for specification and estimation of the model.

Area Models and Forecasts--Estimated models for Areas 2, 3, and 4 are reported in Table 6. Actual observations and model forecasts are plotted in Figure 5. The coefficient estimates and summary statistics from the model for Regulatory Area 2 are shown in the second column of Table 3. For comparison, the sample mean over the 45 observations used for specification and estimation is 94.8 million pounds and the sample standard deviation is 5.7 million pounds. A naive model that predicts the sample average for the eight out-of-sample years produces a root-mean-squared error of 18.9 million pounds, in contrast to the model root-mean-squared error of 4.5 million pounds.

The third column of Table 6 presents the coefficient estimates and summary statistics for IPHC Regulatory Area 3. The in-sample mean for this area is 142.1 million pounds and the standard deviation is 7.0 million pounds. The out-of-sample root-mean-squared error of 3.6 million pounds is far smaller than that of the naive model (forecasting the sample mean), 34.1 million pounds.

Area 4 estimates are given in the fourth column of Table 6. The comparable in-sample mean is 22.5 million pounds, and the sample standard deviation is 2.8 million pounds. The naive model out-of-sample forecast root-mean-squared error is 13.5 million pounds in contrast to the corresponding model statistic of 1.3 million pounds.

Subarea Models and Forecasts--In practice, the IPHC sets catch limits by regulatory subarea.. Table 7 presents univariate models representing regulatory subareas 2A, 2B, and 2C. The corresponding statistics for Regulatory Area 3 subareas 3A and 3B are given in Table 8. Plots of actual observations, model predictions, and errors for each subarea are graphed in Figure 6.

Model Analysis--One conclusion from the summary statistics and figures presented above is that all of the models forecast surprisingly well. Area 2A presents some minor difficulty because two of the errors are large relative to the atypically small mean, but even these forecasts are usable despite the small out-of-sample correlation between forecast and actual values. In the other cases there is no real deterioration out-of-sample, and the errors are minor, apart from area 2A, the smallest out-of-sample correlation is 90%, and many are 99%. At a technical level, the model tits

are so close that these models are unusual in another respect. Unreported investigations show that extending the lag on the autocorrelations included in the Hankel matrix does not significantly change the model; the second-order model tracks autocorrelations at longer lags as well as it tracks the them over the first three lags. Thus we can base the models on the short lags presented without loss of dynamic detail.

While the biomass plots differ markedly at the subarea level, with the exception of subarea 2A, the estimated models are quite similar both to each other and to the more aggregated models.[6] For example, the eigenvalues of the dynamics matrix A are 0.888 ± 0.169i for Regulatory Area 2, and 0.900 ± 0.166i and 0.889 ± 0.175i for subareas 2B and 2C, respectively. The moduli associated with these values are 0.904, 0.915, and 0.906, respectively. Similarly, the Regulatory Area 3 eigenvalues are 0.884 ± 0.216i with values of 0.872 ± 0.222i and 0.911 ± 0.197i for subareas 3A and 3B. The moduli are 0.910, 0.900, and 0.932, respectively. In Regulatory Area 4 the eigenvalues of A are 0.902 ± 0.173i with modulus 0.918. The other coefficients in these models are similar as well. This suggests that while different initial conditions and random exogenous shocks affect the separate subareas, they all share a common dynamic specification. At any given time a particular age distribution of fish is determined by past catch and other factors. Any change in biomass is a reaction of this long-lived slow-to-mature stock to endogenous and exogenous forces (including catch); this reaction is highly predictable using the models presented above. The off-shifting of periods of high and low biomass across regions reflects local shocks rather than stock diversity. Regardless of the source of variation, it is optimal to manage the stocks at as fine a level of disaggregation as is possible given the measurement apparatus.

[6]The eigenvalues for subarea 2A are 0.774 ± 0.118i with modulus 0.783, substantially different from the corresponding values for other areas. Subarea 2A is the southernmost region, an area extending from the Canadian-U. S. border **southward** and is believed to have been fished excessively at the turn of the century. Thus it may be different because of its fringe location, or perhaps because it has been most heavily fished.

Table 5. --A system-theoretic model of coastwide Pacific halibut biomass.

	A		B	C	
Coefficients	0.884	0.190	1.383	0.983	-0.093
	-0.190	0.903	1.600		
Summary statistics	In-sample		Out-of-sample		
Mean (million lbs)	259.52		217.97		
Root-mean-squared error	9.382		6.679		
Percent root-mean-squared error	3.615		3.064		
Average error	0.065		-1.400		
Mean absolute deviation	7.665		6.069		
Correlation	0.99		0.99		

Table 6. -- System-theoretic models of Pacific halibut biomass: IPHC Regulatory Areas 2, 3, and 4.

	Area 2		Area 3		Area 4	
Coefficients						
A	0.902	0.170	0.858	0.217	0.881	0.174
	-0.170	0.875	-0.217	0.910	-0.174	0.924
B	1.391	1.755	1.485	1.757	1.185	0.984
C	0.985	-0.085	0.982	-0.104	0.973	-0.083
In-sample statistics						
Mean (million lbs)	94.841		142.14		22.543	
Root-mean-squared error	3.247		5.807		1.140	
Percent root-mean-squared error	3.424		4.085		5.057	
Average error	0.010		0.093		-0.047	
Mean absolute deviation	2.606		4.300		0.949	
Correlation	0.99		0.99		0.99	
Out-of-sample statistics						
Mean (million lbs)	77.638		130.98		9.351	
Root-mean-squared error	4.495		3.637		1.326	
Percent root-mean-squared error	5.790		2.777		14.180	
Average error	-2.064		1.128		-1.016	
Mean absolute deviation	3.366		3.064		1.086	
Correlation	0.91		0.99		0.96	

Table 7. -- System-theoretic models of Pacific halibut biomass: IPHC Regulatory Subareas 2A, 2B, and 2C

	Area 2A		Area 2B		Area 2C	
Coefficients						
A	0.875	0.155	0.897	0.166	0.905	0.175
	-0.155	0.673	-0.166	0.902	-0.175	0.873
B′	1.262	1.244	1.262	1.266	1.655	2.632
C	0.972	-0.084	0.980	-0.081	0.990	-0.089
In-sample statistics						
Mean (million lbs)	2.801		51.776		40.264	
Root-mean-squared error	0.162		2.254		1.328	
Percent root-mean-squared error	5.784		4.353		3.298	
Average error	-0.003		-0.042		0.058	
Mean absolute deviation	0.133		1.824		1.054	
Correlation	0.99		0.99		0.99	
Out-of-sample statistics						
Mean (million lbs)	1.019		31.669		44.950	
Root-mean-squared error	0.235		2.008		2.837	
Percent root-mean-squared error	23.062		6.341		6.311	
Average error	-0.204		-1.648		-0.777	
Mean absolute deviation	0.204		1.789		2.157	
Correlation	0.10		0.93		0.90	

Table 8. -- System-theoretic models of Pacific halibut biomass: IPHC Regulatory Subareas 3A and 3B.

	Area 3A		Area 3B	
Coefficients				
A	0.853	0.222	0.869	0.201
	-0.222	0.892	-0.201	0.952
B′	1.557	1.976	1.349	1.283
C	0.983	-0.108	0.980	-0.095
In-sample statistics				
Mean (million lbs)	98.306		43.833	
Root-mean-squared error	4.003		1.954	
Percent root-mean-squared error	4.072		4.458	
Average error	0.055		0.081	
Mean absolute deviation	2.935		1.570	
Correlation	0.99		0.99	
Out-of-sample statistics				
Mean (million lbs)	95.893		35.085	
Root-mean-squared error	4.118		1.308	
Percent root-mean-squared error	4.294		3.728	
Average error	1.750		-0.840	
Mean absolute deviation	3.696		1.041	
Correlation	0.99		0.99	

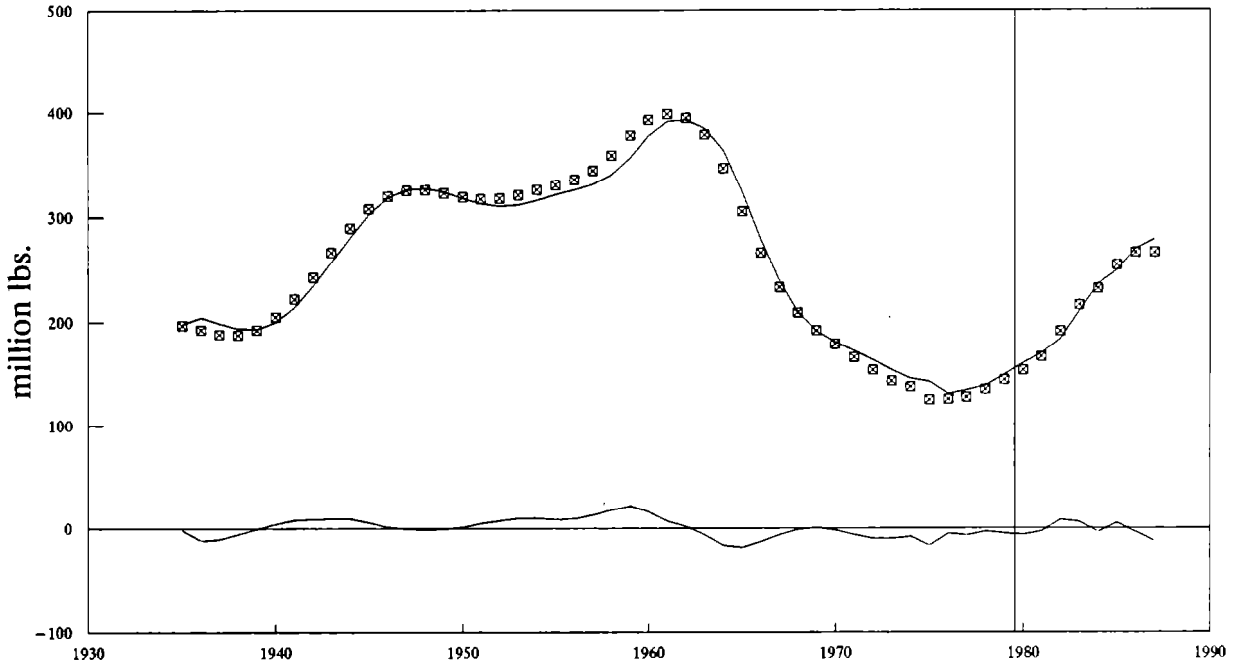

Figure 4. --A system-theoretic time-series model of coastwide Pacific halibut biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are Represented by solid lines.

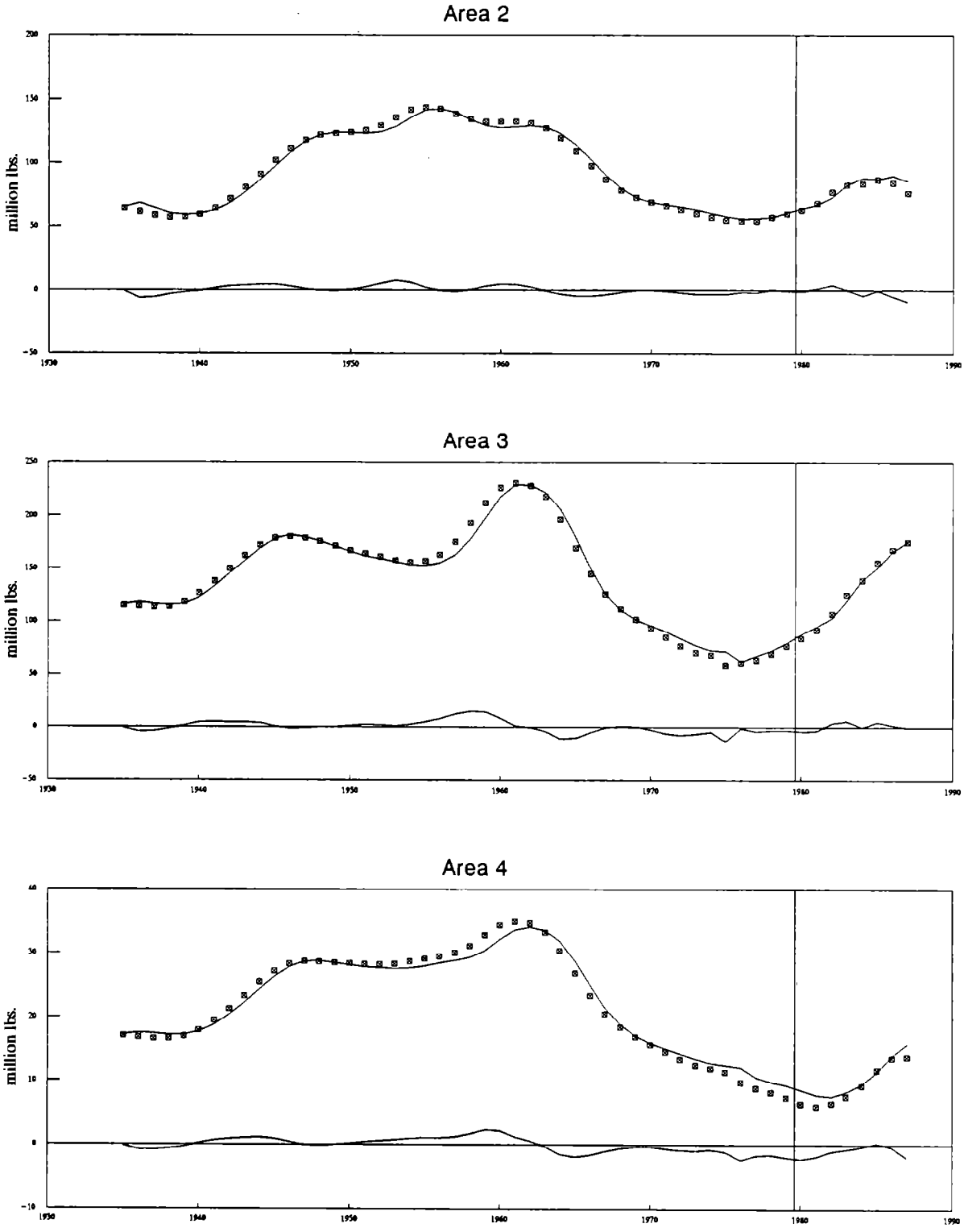

Figure 5.-- System-theoretic time-series models of Pacific halibut biomass: IPHC Regulatory Areas 2, 3, and 4. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

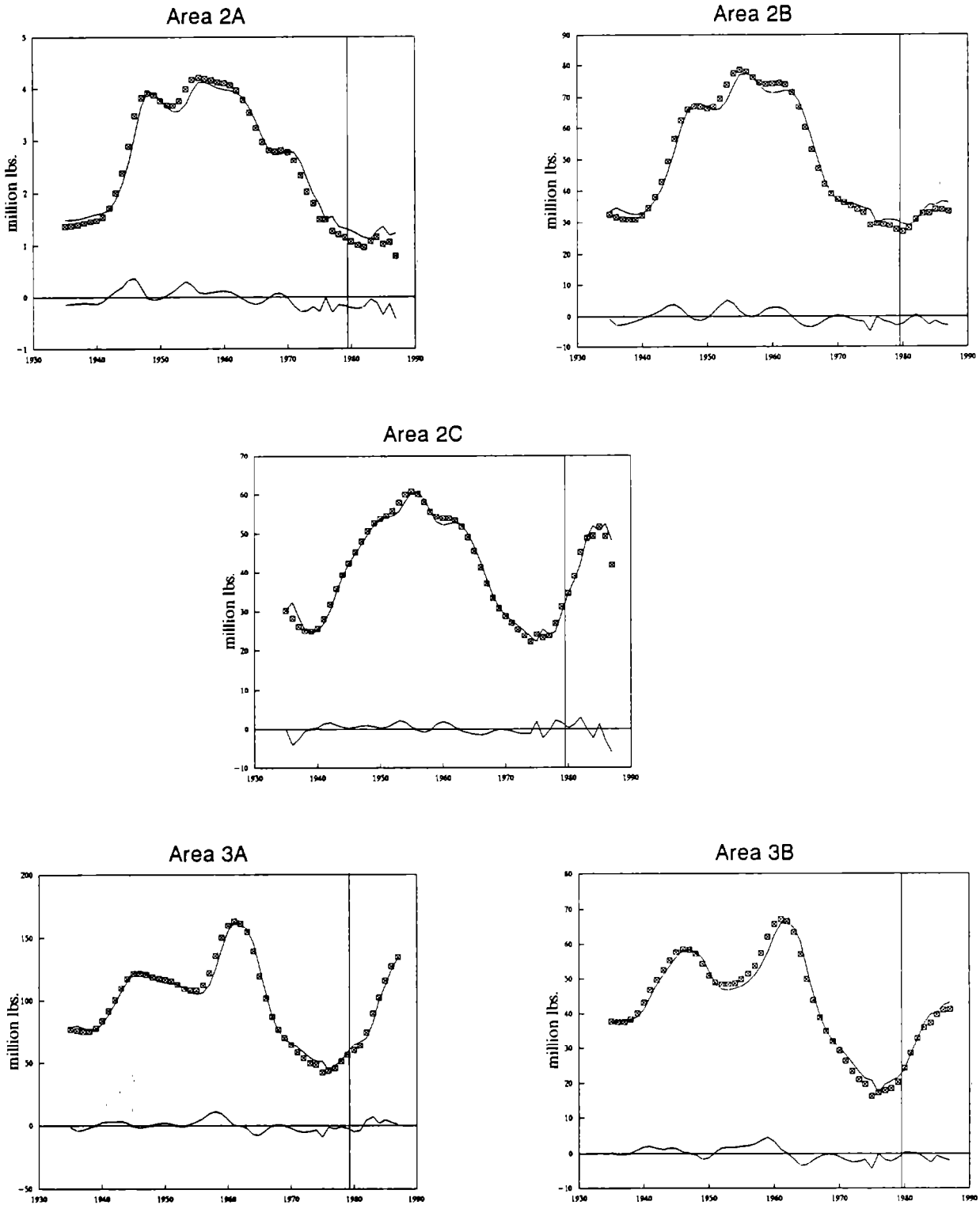

Figure 6. --System-theoretic time-series models of Pacific halibut biomass: IPHC Regulatory Subareas 2A, 28, 2C, 3A, and 38. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

APPROXIMATE STRUCTURAL/TIME-SERIES MODELING

Although the true form of population dynamics is unknown, approximate encompassing structural models can be used to produce forecasts with stationary residuals with a complex and perhaps; nonlinear serial correlation structure. Time-series analysis can be applied to these residuals to improve forecast accuracy. Because even nonlinear processes have linear state-space representations so long as they are stationary, the system-theoretic time-series model is particularly well suited for determining higher order dynamics from the residuals of the approximate model.

Approximate Structural/System-theoretic Time-series Models

Combining the approximate structural model with the system-theoretic time-series methodology results in the approximate encompassing system-theoretic model:

$$x_t = \beta_0 + \beta_1 x_{t-1} - h_{t-1} + \varepsilon_t , \tag{9}$$

$$D\varepsilon_t = DCz_{t|t-1} + De_t , \text{ and} \tag{11}$$

$$z_{t+1|t} = Az_{t|t-1} + [BD^{-1}]De_t . \tag{12}$$

Equation (9) represents the approximate structure of the system, and may include lagged endogenous variables (x_{t-i}), as well as current and lagged predetermined variables. Equation 11 and Equation 12 represent the higher order dynamics and nonlinearities of the residuals (ε_t).

It is assumed that ε_t is normally distributed and contemporaneously correlated, that is, $\varepsilon_t \sim N(\mu_\varepsilon, \Delta)$. Efficient estimation requires that the error-covariance matrix be taken into consideration in the estimation of the coefficients of the structural model. The error-covariance matrix A is unknown, but it can be estimated conditional on B_i, and is of the form:

$$\Delta = \begin{bmatrix} \Gamma_0 & DC\Omega & DCA\Omega & \dots & DCA^{T-1}\Omega \\ DC\Omega & \Gamma_0 & DC\Omega & & DCA^{T-2}\Omega \\ DCA\Omega & DC\Omega & \Gamma_0 & & . \\ . & & & & \\ DCA^{T-1}\Omega & DCA^{T-2}\Omega & \dots & DC\Omega & \Gamma_0 \end{bmatrix},$$

where Γ_0 and Ω are defined by Equations (14) and (18), respectively. Because the estimated covariance matrix is conditional on the coefficients B_i, and the coefficient estimates are conditional on A, the model must be iterated to convergence.

Given estimates of the Bi and A, Equations (11) and (12) can be solved for estimates of A, C, and B as shown earlier. With estimates of A, B, C, and B, and a set of initial conditions z_o, Equations (9), (11), and (12) can be solved for in- and out-of-sample forecasts.

Pacific Halibut [7]

In addition to the biomass data used in the system-theoretic time-series models of Pacific halibut biomass, the models in this section also incorporate data on landings. Commercial setline catches for 1935-84 are reported in Quinn et al. (1985, Appendix Table 10). More recent setline catches are reported in the 1987 Annual Report of the International Pacific Halibut Commission (1988, Table 2). Model specification and parameter estimation are based on 48 observations from 1935 to 1982, reserving the 5 observations from 1983 to 1987 for out-of-sample validation. Four models were estimated: 1) a first-order linear approximate model without time series (OLS); 2) a Schaefer-type logistic surplus yield model; 3) a Ricker-type delay-difference model with logistic growth and recruitment functions and an 8-year recruitment lag; and 4) a first-order linear state-space model iterated to convergence (SITS). Estimated coefficients are reported in Table 9.

Both coefficients of the linear and quadratic first-order approximate models are significantly different from zero with confidence levels in excess of 99%. Although the delay-difference model outperformed the simple linear and logistic models on many measures of performance (see below), only the linear coefficient is significantly different from zero. While both structural coefficients of the STTS model are significantly different from zero at a 95% confidence level, the intercept term is only just significant, suggesting that omitted higher-degree terms are also of-low significance.

[7]The following example is reported in Criddle and Havenner (1991).

Model Validation and Forecasts--Performance of the four models is compared in Table 10 and Figure 7. Actual observations are plotted as crosses, predictions are represented as a solid line through the data, and errors are shown as a solid line about zero. Reference lines are included at zero biomass and at the end of data used for model specification and estimation.

All four models perform well in the sample estimation portion of the data set. The Schaefer and OLS models are virtually indistinguishable, suggesting that nonlinearities are not significant. The more general Ricker-type model reduces root-mean-squared error (RMSE) by about 1% or 1.5 million pounds. Because the Ricker model was specified to embed the Schaefer model, all error reductions can be attributed to the specification of stock-recruit dynamics. Nevertheless, caution should be exercised in extending favor to the Ricker model because the hypothesis that the restrictions $f(\mathbf{x}_{t-i}) = 0$ for i = 2,3,4,5,6,7 is rejected at the 1% level ($F_{12,24} = 6.192$), and because the estimated stock-recruit coefficients are not significant (see Table 9). The state-space model reduces RMSE an additional 1.6 million pounds (-1%) below the RMSE of the Ricker model. The portmanteau statistic indicates that the residuals of the STTS model are not serially correlated.

Because the models differ in the number of coefficients to be estimated and in the number of observations required for initialization, performance was also contrasted on the basis of Schwartz's criterion.[8] The relative ranking of the models is the same as under the RMSE criterion!: STTS, Ricker, Schaefer, and OLS.

Out-of-sample validation was less definitive. The percentage root-mean-squared errors (%RMSE) of all four models declined in the out-of-sample portion of the data set. The Ricker and STTS models continued to dominate the other two models. However, the Ricker model generated forecasts with lower percentage root-mean-squared errors than the STTS model, which also had the smallest mean absolute deviation.

Lehmann (1959) suggests a test for the equality of mean-squared forecast errors. Estimated correlation coefficients are reported in Table 11, with correlations that are significantly

[8]The Ricker model required eight observations for initialization. All other models required one. The number of coefficients estimated were two, two, four, and nine for the OLS, Schaefer, Ricker. and STTS models, respectively.

different from zero at a 90% confidence level denoted by an asterisk. In-sample correlations clearly indicate that the mean-squared error of the STTS model is significantly different from that of the Ricker model, which in turn differs significantly from the root-mean-squared errors of the Schaefer and OLS models, while the residuals cannot be distinguished from each other. The out-of-sample correlations do not distinguish the Ricker model from any of the other models, while the mean-squared error of the STTS model is significantly smaller than those of the Schaefer and OLS models.

Model Stability--Static sustainable yields can be determined for the four models given any realization of biomass. Static sustainable yields are those harvest levels that result in constant biomass; that is, h_t such that $x_t = x_{t-1}$ for all t.

Comparison of the set of sustained yields corresponding to the time-series of coastwide biomass with actual harvest levels reveals some interesting patterns. All four models indicate that actual harvests exceeded sustainable harvests for every year from 1959 through 1972. Over this same period, coastwide biomass varied from a high of 307 million pounds in 1959 to a low of 134 million pounds in 1972. Although setline landings were consistently below sustainable levels from 1973 to 1985, biomass continued to drop through 1975 before responding to reduced harvesting. Taken together, these observations suggest that the biomass decline which occurred in the 1960s and early 1970s was a direct response to elevated setline landings, while recent increases in abundance are due to reduced landings. However, even when catches were reduced below sustainable levels, there was a 3-year lag in the response observed in the biomass. Landings again exceeded sustainable harvest levels in 1986 and 1987, so another decline in coastwide abundance is likely.

Sustainable and actual landings are compared in Table 12. Actual setline landings have averaged about 1 million pounds less than sustainable harvests, However, the root-mean-squared difference between actual and sustainable harvest levels has been near 20% of mean landings. Figure 8 plots actual setline landings, sustainable harvest levels, and the difference between sustainable and actual catches.

Table 9. -- Approximate, structural, and approximate structural/time-series models of Pacific halibut biomass: Coefficients.

OLS

$$x_t = \underset{(3.67)}{16.763} + \underset{(57.92)}{1.162}x_{t-1} - h_{t-1} + \varepsilon_t$$

Schaefer

$$x_t = \underset{(53.31)}{1.344}x_{t-1} - \underset{(-4.54)}{.0004}x_{t-1}^2 - h_{t-1} + \varepsilon_t$$

Ricker

$$x_t = \underset{(11.01)}{1.4261}x_{t-1} - \underset{(-1.79)}{.0005}x_{t-1}^2 - \underset{(-.78)}{.0963}x_{t-8} + \underset{(.41)}{.0001}x_{t-8}^2 - h_{t-1} + \varepsilon_t$$

STTS

$$x_t = \underset{(1.84)}{19.018} + \underset{(25.03)}{1.156}x_{t-1} - h_{t-1} + \varepsilon_t$$

$$\varepsilon_t = \begin{bmatrix} -7.896 & 1.626 \end{bmatrix} z_{t|t-1} + e_t$$

$$z_{t+1|t} = \begin{bmatrix} .751 & .178 \\ -.178 & .306 \end{bmatrix} z_{t|t-1} + \begin{bmatrix} -.106 \\ -.088 \end{bmatrix} e_t$$

x_t is the coastwide biomass of Pacific halibut measured in million pound units. T-statistics are reported in parentheses below the coefficients.

Table 10. -- Approximate, structural, and approximate structural/time-series models of Pacific halibut biomass: Statistics.

	Mean	%RMSE	R^2	MAD	Q	SC
In-sample (48 obs)						
OLS	218.91	4.02	0.981	7.587	98.36	4.513
Schaefer	218.91	3.79	0.983	7.153	95.79	4.397
Ricker	224.40	3.08	0.990	5.803	121.06	4.238
STTS	218.91	2.41	0.993	4.308	19.40	3.981
Out-of-sample (5 obs)						
OLS	246.65	3.58	0.799	8.077		
Schaefer	246.65	3.07	0.852	6.528		
Ricker	246.65	1.74	0.952	3.910		
STTS	246.65	1.98	0.938	3.514		

%RMSE is, root-mean-squared error as a Percent of mean biomass. MAD is mean absolute deviation in units of one millions pounds; Q is the Ljung-Box portmanteau test statistic and is distributed asymptotically x^2 with 18, 18, 16, and 11 degrees of freedom for the OLS, Schaefer, Ricker, and STTS models, respectively; SC is Schwartz's criterion; and Q and SC are not reported for the out-of-sample forecasts due to the small number of observations.

Table 11. --. Approximate, structural, and approximate structural/time-series models of Pacific halibut biomass: Error correlations.

	STTS	OLS	Schaefer	Ricker
In-sample (48 obs)				
STTS	1.000	-0.451*	-0.419*	-0.296*
OLS	0.451*	1.000	0.147	-0.301*
Schaefer	0.419*	-0.147	1.000	0.252
Ricker	0.296*	-0.301*	-0.252	1.000
Out-of-sample (5 obs)				
STTS	1.000	-0.639	-0.530	0.248
OLS	0.639	1.000	0.761	0.856*
Schaefer	0.530	-0.761	1.000	0.846*
Ricker	-0.248	-0.856*	-0.846*	1.000

Correlations that are significantly different from zero at a 90% confidence level are superscripted with an asterisk.

Table 12. -- Approximate, structural, and approximate structural/time-series models of Pacific halibut biomass: Sustainable yields.

	Mean	AVERR	%RMSE	CORR	MAD
OLS	51.42	1.05	18.3	0.79	7.83
Schaefer	51.42	1.39	18.1	0.80	7.58
Ricker	51.62	0.67	22.5	0.70	9.64
STTS	51.42	1.96	18.7	0.79	7.85

Mean is the average of annual landings in 1 million lb. units; AVERR is the mean difference between the steady state harvest level and actual landings; %RMSE is root-mean-squared error as a percent of mean landings; and MAD is mean absolute deviation.

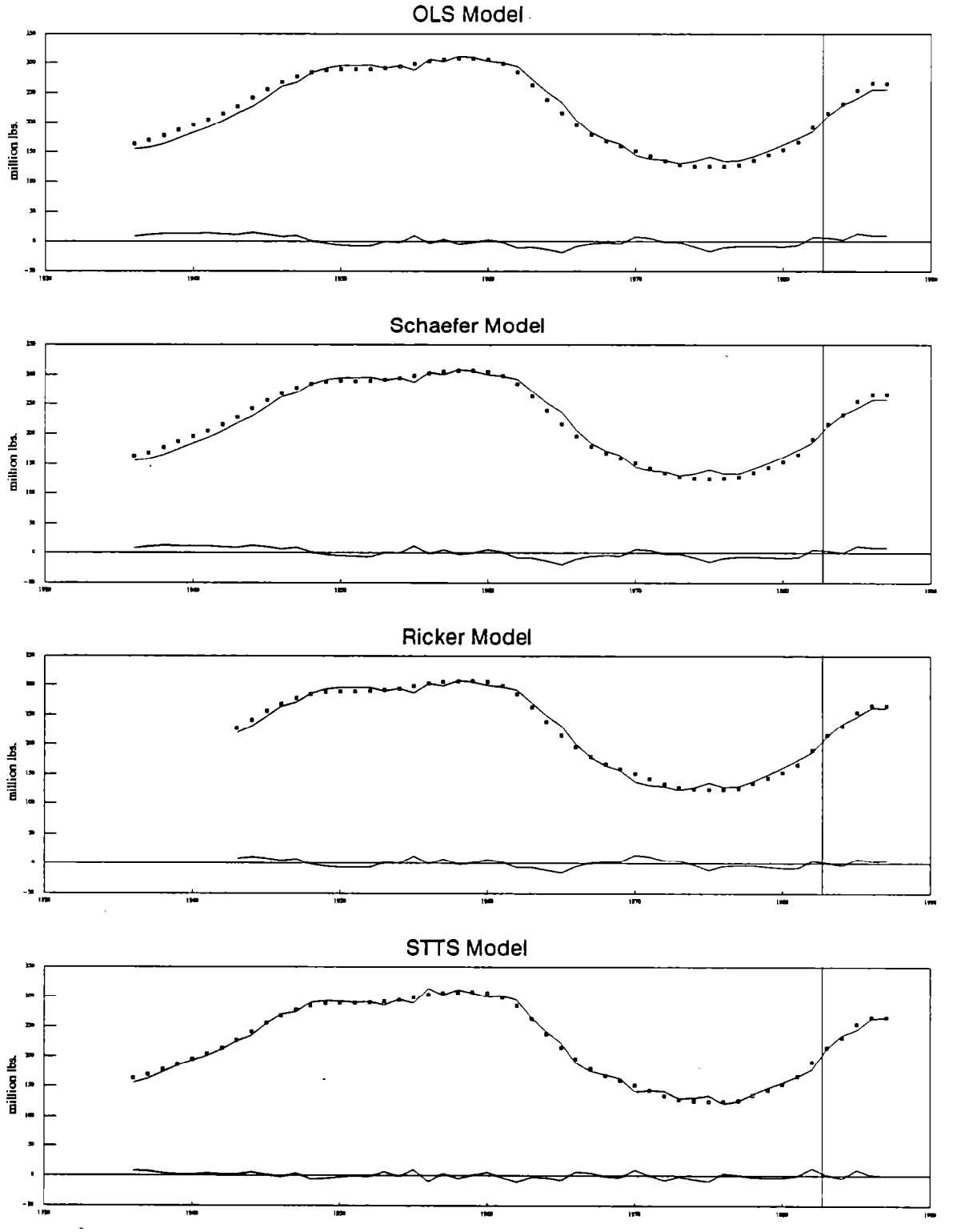

Figure 7. -- Forecasts from structural, time-series, and approximate structural/time-series models of coastwide Pacific halibut biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

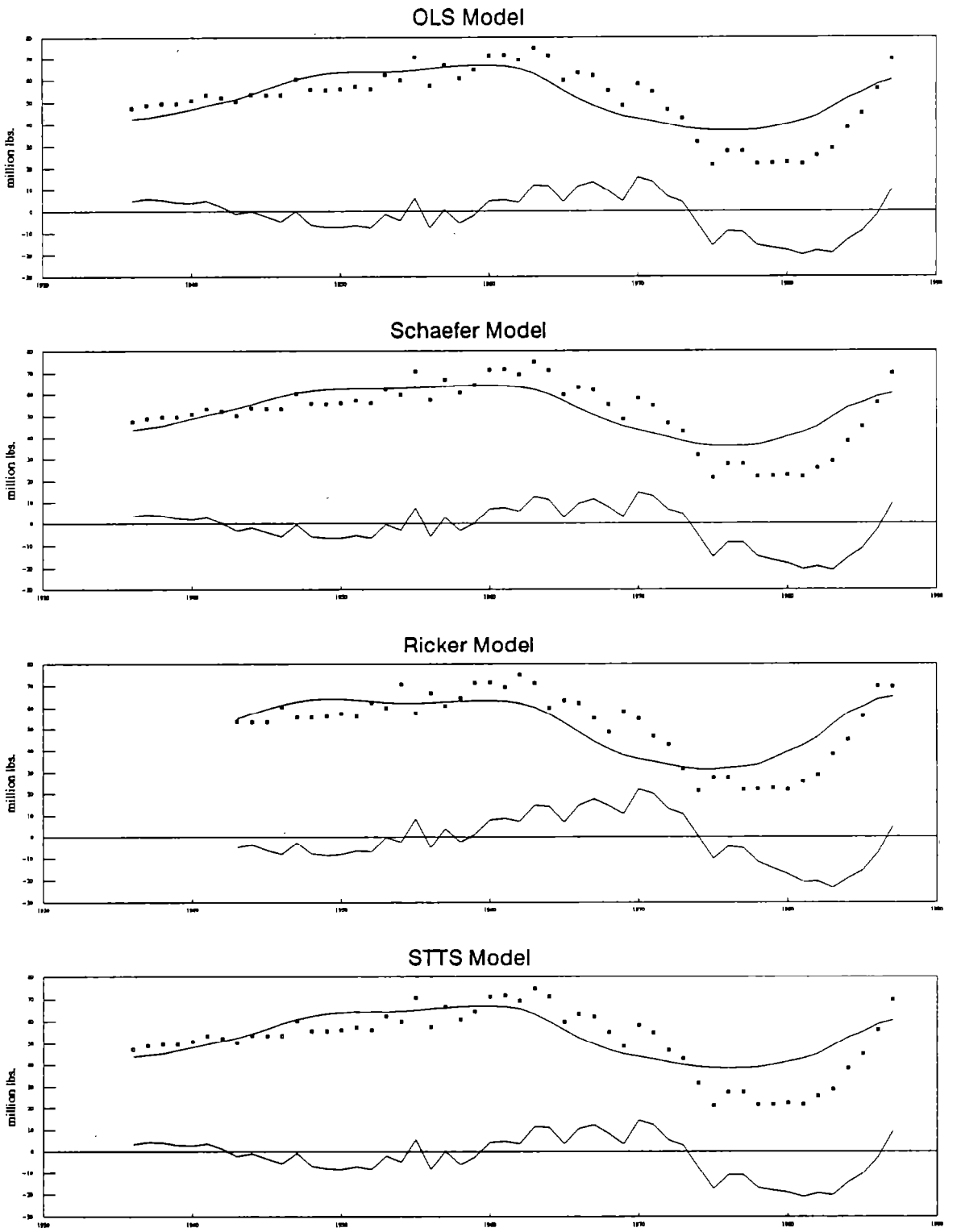

Figure 8. --Estimates of the coastwide sustainable yield of Pacific halibut derived from structural, time-series, and approximate structural/time-series models. Actual landings are denoted by crossed boxes. Estimates of sustainable yield and excess harvests are represented by solid lines.

Pacific Halibut, Yellowfin Sole, and Red King Crab

This section will extend the restricted reduced-form approximate encompassing model of Pacific halibut, yellowfin sole, and red king crab biomass developed above by applying system-theoretic time-series to the residuals of the maximum likelihood (MLE) and constrained indirect least-squares (CIDLS) first-order linear models. Coefficient estimation will be based on the 18 annual (observations of catch and biomass from 1969 through 1986 (Table 3). Performance of the maximum likelihood and constrained indirect least squares models will be compared to models which apply system-theoretic time-series to the residuals of corresponding approximate structural models. An additional model based on STTS without a structural filter is also presented. Coefficient estimates for the five models are reported in Table 13.

Model Performance--The three time-series models each required a five-period system lag. Because the restricted approximate encompassing time-series models imposed first-order dynamics in their approximations to the encompassing model, the residuals had a system lag of four periods.

Summary statistics for the five models are included as Table 14. The unrestricted approximate encompassing system-theoretic time-series model is preferred over all of the other models on the basis of every criteria except log likelihood and Schwartz's criterion. The $CIDLS_{STTS}$ model is nearly as successful as the MLE_{STTS} model, and has the highest log-likelihood value. The approximate models without time-series have the smallest Schwartz's criterion values as a result of the smaller number of coefficients specified. Nine coefficients were estimated for the MLE model, 6 for the CIDLS model, 34 for the pure time-series model, and 36 and 26 for the joint MLE_{STTS} and $CIDLS_{STTS}$ models, respectively. The STTS model was least successful, particularly on the king crab series. Performance of the five models is easily compared in plots of observed biomass, model predictions, and residuals. Figures 9, 10, and 11 include estimated biomass, predictions, and errors generated by the five models for halibut, yellowfin sole, and red king crab, respectively.

Table 14 also includes 1987 biomass forecasts for each model based on information available in 1986. The actual 1987 estimated biomass was 5,764 t for Pacific halibut and

20,827 t for red king crab. Cohort analysis estimates for yellowfin sole were not available.

Performance of all four linear first order approximate models is good. By contrast, the pure time-series model is less successful. In light of the successes achieved with system-theoretic methods, the poor performance of this STTS model is somewhat surprising. The explanation underlines the importance of matching model design and estimation techniques to available data. Because of the shortness and volatility of the data set, STTS is an inappropriate technique in this application. Over the period from 1969 through 1986, Pacific halibut biomass declined; yellowfin sole biomass increased; and the biomass of red king crab varied from 12,290 t to 127,710 t and then down to 5.910 t. Over this short time-series, the data violate the stationarity requirement for estimation using pure time-series methods. Approximations to the encompassing model explain observed trends and have stationary residuals.

When the data set used for estimation is more extensive, time-series techniques will have improved performance. Table 6 reports coefficient estimates and summary statistics for IPHC Regulatory Area 4. Because the sample used for estimation is stationary, the dynamics of the estimated model are simple, and the %RMSE value of 5.1 is smaller than the corresponding value for multivariate estimates of halibut biomass. It is unlikely that multispecies modeling results in improved forecasts of halibut biomass because information regarding the abundance of the other species modeled limits the number of observations of halibut abundance that can be used for estimation. As the time-series of observations on other species lengthens, gains to multispecies modeling are also likely to occur.

Multispecies modeling results in improved accuracy in the estimation of yellowfin sole biomass. Table 2 reports the performance of univariate linear, quadratic, and cubic approximate models of yellowfin sole biomass: All four multivariate linear approximate models had lower percentage root-mean-squared errors than the linear univariate model, and the MLE_{STTS} had a lower %RMSE than any of the univariate yellowfin sole models.

Table 13. -- Pacific halibut, yellowfin sole, and red king crab biomass: Coefficients.

Maximum Likelihood

$$x_t = \begin{bmatrix} 1.01 & 0.81 & -0.01 \\ 0.02 & 1.03 & 0.001 \\ 3.41 & -8.95 & 1.28 \end{bmatrix} x_{t-1} - h_{t-1} + \varepsilon_t$$

Constrained Indirect Least Squares

$$x_t = \begin{bmatrix} 1.04 & 0.75 & -0.01 \\ & 1.07 & 0.002 \\ & & 1.32 \end{bmatrix} x_{t-1} - h_{t-1} + \varepsilon_t$$

System-theoretic Time-series

$$x_t = \begin{bmatrix} 0.71 & -0.73 & 0.17 & 0.62 \\ -0.53 & -0.02 & -0.19 & 0.31 \\ 8.51 & 37.75 & 6.22 & 0.35 \end{bmatrix} z_{t|t-1} + e_t$$

$$z_{t+1|t} = \begin{bmatrix} 0.94 & -0.34 & -0.07 & -0.05 \\ 0.37 & 0.77 & -0.11 & 0.09 \\ 0.01 & -0.01 & 0.89 & 0.71 \\ 0.04 & -0.14 & -0.31 & 0.66 \end{bmatrix} z_{t|t-1} + \begin{bmatrix} 0.23 & 0.69 & 0.01 \\ -0.13 & 2.10 & 0.01 \\ 1.02 & -1.75 & 0.02 \\ 0.29 & 2.63 & 0.01 \end{bmatrix} e_t$$

System-theoretic Time-series Model of Maximum Likelihood Estimate Residuals

$$\varepsilon_t = \begin{bmatrix} -0.24 & -0.05 & -0.07 \\ 0.05 & 0.004 & 0.02 \\ -3.48 & -9.47 & -0.70 \end{bmatrix} z_{t|t-1} + e_t$$

$$z_{t+1|t} = \begin{bmatrix} 0.45 & 0.71 & -0.34 \\ -0.64 & 0.58 & 0.06 \\ -0.14 & 0.21 & 0.67 \end{bmatrix} z_{t|t-1} + \begin{bmatrix} -0.04 & 4.50 & -0.02 \\ 0.28 & -4.97 & -0.01 \\ -1.22 & 1.11 & 0.01 \end{bmatrix} e_t$$

System-theoretic Time-series Model of Constrained Indirect Least-Squares Residuals

$$\varepsilon_t = \begin{bmatrix} -0.03 & -0.23 \\ -0.01 & 0.04 \\ -12.79 & -8.17 \end{bmatrix} z_{t|t-1} + e_t$$

$$z_{t+1|t} = \begin{bmatrix} 0.59 & -0.60 \\ 0.69 & 0.46 \end{bmatrix} z_{t|t-1} + \begin{bmatrix} -0.20 & -4.71 & -0.01 \\ 0.31 & 3.83 & -0.001 \end{bmatrix} e_t$$

The vector x_t is ordered: Pacific halibut, yellowfin sole, and red king crab. Pacific halibut and red king crab biomass is in thousands of metric tons. Yellowfin sole biomass is in millions of metric tons.

Table 14. -- Pacific halibut, yellowfin sole, and red king crab biomass: Statistics.

	MLE	CIDLS	STTS	MLE_{STTS}	$CIDLS_{STTS}$	
R^2						
Halibut	0.922	0.917	0.847	0.954	0.953	
Sole	0.979	0.973	0.960	0.982	0.978	
King crab	0.853	0.800	0.663	0.928	0.920	
Percentage Root-Mean-Squared Error						
Halibut	7.1	7.3	9.9	5.4	5.5	
Sole	8.2	9.3	11.3	7.6	8.5	
King Crab	35.5	41.5	53.9	24.9	26.3	
Mean Absolute Deviation						
Halibut	0.240	0.270	0.340	0.176	0.175	
Sole	0.069	0.080	0.106	0.069	0.073	
King Crab	12.81	14.354	18.371	8.383	10.135	
Ljung-Box Q						
Halibut	12.49	12.47	12.90	9.01	6.79	
Sole	9.29	7.03	8.32	5.30	4.85	
King Crab	13.91	24.28	7.41	5.75	6.10	
Log Likelihood	-59.6	-63.6	-73.7	-45.4	-45.2	
Schwartz's Criterion	155.2	151.0	283.0	234.3	194.1	
	Out-of-sample forecasts (1987)					Actual
Halibut	6.111	6.174	5.742	6.038	6.059	5.764
Sole	1.762	1.727	1.336	1.772	1.744	
King Crab	18.314	13.925	56.260	8.573	3.197	20.827

Halibut and crab are reported in thousand metric ton units; yellowfin sole is reported in million metric tons.

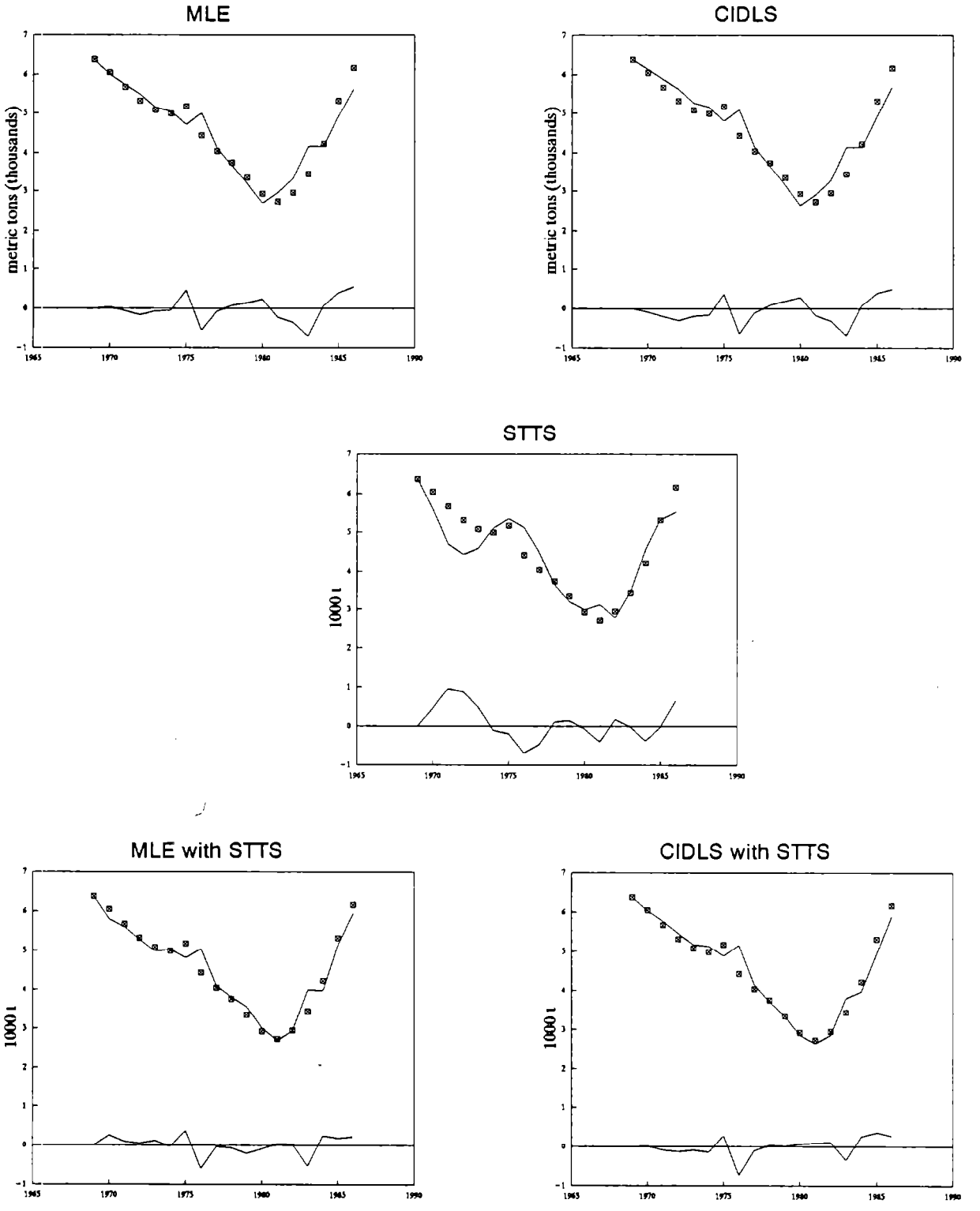

Figure 9. -- Five models of Pacific halibut biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

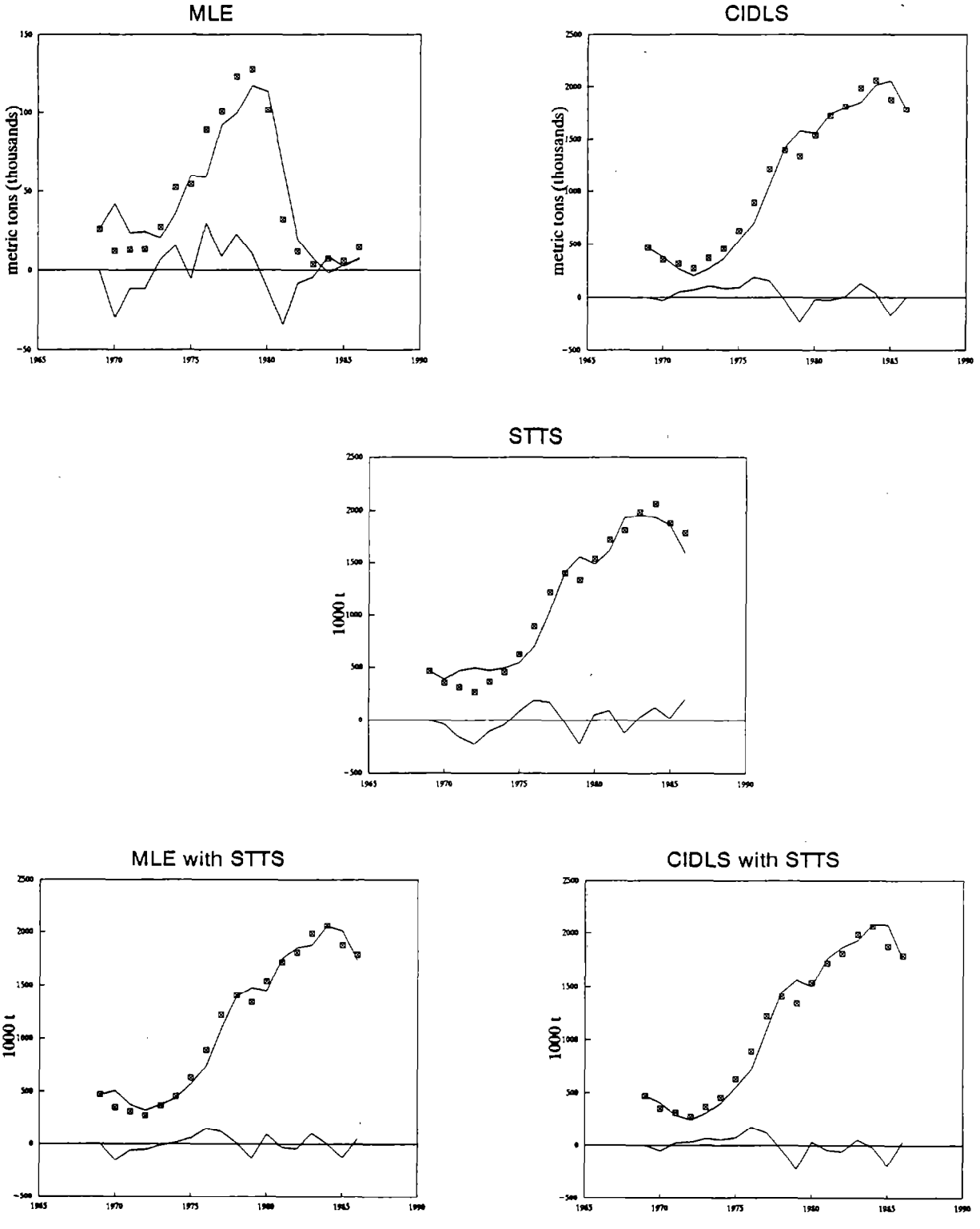

Figure 10 .-- Five models of yellowfin sole biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

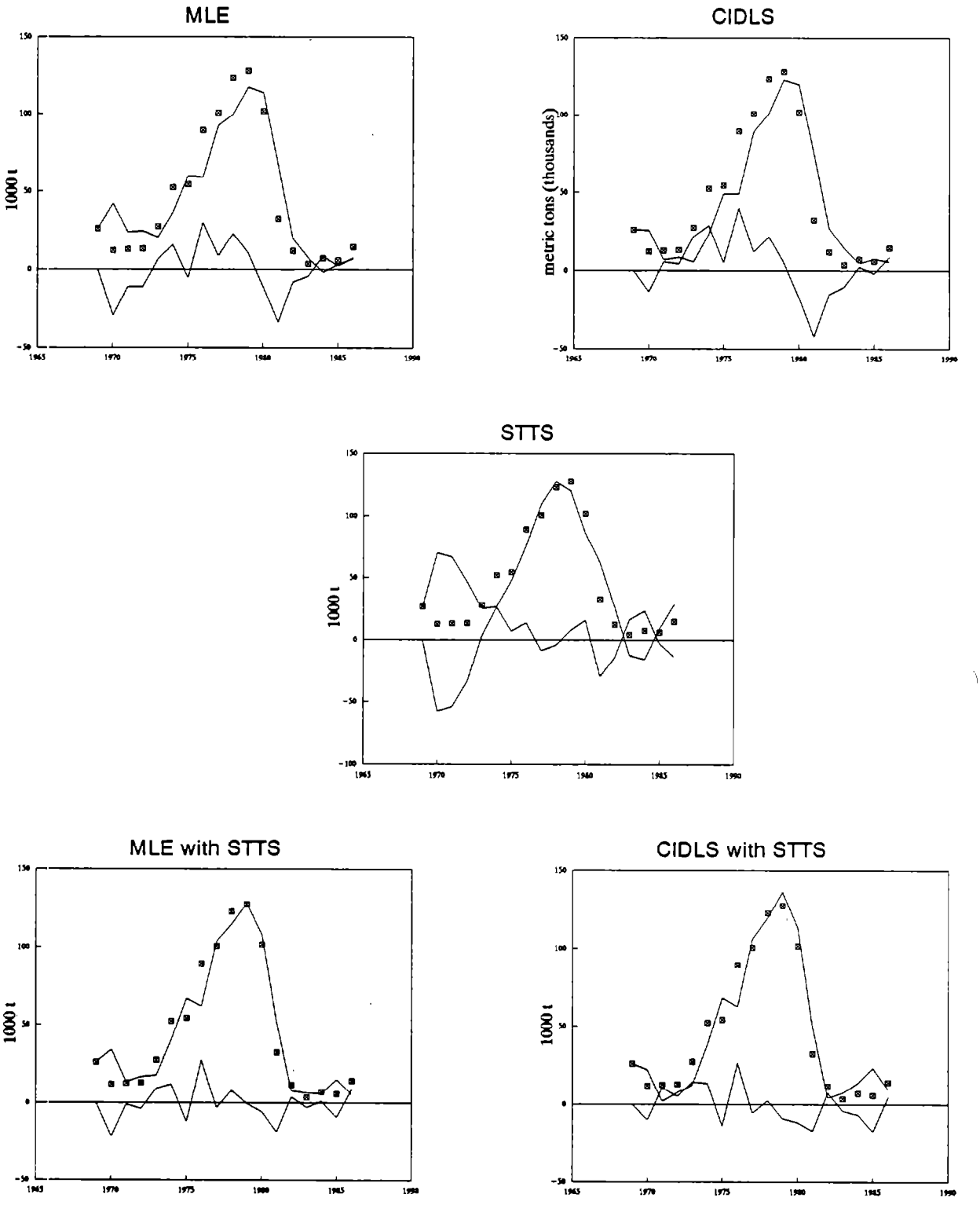

Figure 11. -- Five models of red king crab biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

Walleye Pollock, Pacific Halibut, Pacific Cod, Yellowfin Sole, and Red King Crab

The preceding example excluded three of the four most abundant target species (groups) of the Bering Sea: walleye pollock (Theragra chalcogramma), Pacific cod (Gadus macrocephalus), and Pacific salmon (Oncorhynchus spp.). A model of the three most important groundfish species plus Pacific halibut and red king crab will be considered in the following model. Pacific salmon will continue to be excluded because they typically do not interact with groundfish and crab. Although there are concerns about the incidental harvest of other species of king (Lithodes spp., Paralithodes spp.) and Tanner crab (Chionoecetes spp.), data regarding their populations are very limited. Catch and abundance data for walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab are reported in Table 15.

Models and Coefficient Estimates--The limited data combined with the long life spans of several of the species makes it difficult to develop models that can be advocated with confidence. The models presented below should be considered as illustrations of the procedure.

All six models presented below are variations on the multispecies first-order linear approximate model. The MLE model allows the fullest range of possible interactions, but requires many coefficients. Restricting the Jacobian matrix reduces the number of coefficients, but also impairs the ability of the models to track the data. Two restricted models are considered: CIDLS, which imposes only those constraints that satisfy a Wald pretest; and seemingly unrelated regression (SUR), which disallows interactions except in the residuals.[9] Three additional models, each applying STTS to the errors of the above models, were also estimated. Coefficient estimates are reported in Table 16, with coefficients significant at a 95% confidence level denoted by an asterisk.

[9]SUR estimates were obtained from

$$\beta_{SUR} = \{[I \otimes vec(X^-)]'[I \otimes vec(X^-)]\}^{-1}[I \otimes vec(X^-)]' vec(X^+),$$

where X^- is a 9 x 5 matrix of normalized biomass from the first through the (n-1)-th time period. Because the series were of very different magnitude, each was divided by its standard deviation to reduce heteroscedasticity. The matrix X^+ is a matrix of normalized biomass from the second through last period plus normalized catch from the first through the (n-l)-th time period. Catch was normalized by dividing each catch series by the standard deviation of the corresponding biomass series.

Structural Stability--The structural dynamics of the MLE and CIDLS models are characterized by stable spirals in two dimensions, unstable spirals in two dimensions, and one dimensional unstable nodes. The SUR model is an unstable node in 5-dimensional space with eigenvalues equal to the diagonal elements of the coefficient matrix.

If stochastics are ignored, each model can be stabilized by specific harvest levels on each species. For example, a deterministic system characterized by the SUR model is in dynamic equilibrium when walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab are harvested at 11, 30, 10, 7, and 7% of their respective biomasses. The average harvest rates over the past 10 years are 13, 27, 11, 6, and 41% for walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab, respectively. Therefore, on the basis of the deterministic SUR model, walleye pollock, Pacific cod, and especially red king crab biomasses are predicted to decline, while biomass of Pacific halibut and yellowfin sole are predicted to increase. These predictions are supported for Pacific halibut and red king crab (see Table 15). The SUR model also suggests that all species were harvested in excess of their deterministic sustainable yields in 1988.[10] Analysis of the stability of the MLE and CIDLS models is more complex because of the interactions among species. All three structural models agree that pollock and red king crab have been harvested in excess of sustainable levels between 1979 and 1988. Both of the structural models with interactions suggest that recent harvests of red king crab have been slightly below sustainable levels, while the SUR model indicates that actual landings continue to exceed sustainable levels by a small amount. Average landings of Pacific halibut were below sustainable levels from 1979 through 1988, but, all three structural models agree that 1987 harvests exceeded sustainable levels, and the SUR model suggests that sustainable levels were also exceeded in 1988.

Results and Discussion--On the basis of R^2, percentage root-mean-squared error, mean absolute deviation, and log likelihood, models were ranked in decreasing order of preference: MLE_{STTS}, MLE, $CIDLS_{STTS}$, CIDLS, SUR_{STTS}, and SUR. Although the performance of each of

[10]Harvesra for 1988 were 19, 37, 21, 8, and 13% of biomass for walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab, respectively.

the structural models improved when the errors were modeled with STTS, the observation set was so short that only one state variable could be chosen to simultaneously explain the errors of all five series. Schwartz's criterion penalizes models with large numbers of coefficients, and results in a different ranking: CIDLS, MLE, $CIDLS_{STTS}$, SUR, MLE_{STTS}, and SUR_{STTS}. All six models were most successful at modeling Pacific halibut biomass and least successful at modeling Pacific cod biomass. Table 17 presents the summary measures of model performance and the model forecasts for 1989.

All models forecast a decline in walleye pollock biomass. Every model, except the SUR model without STTS treatment of the error vector, forecasts a slight increase in the biomass of Pacific halibut. The four SUR and MLE models predict declines in cod biomass, while the two CIDLS models predict slight increases. Even though all six models underestimated 1988 yellowfin sole biomass, the four MLE and CIDLS models predict substantial future declines in yellowfin sole biomass. The SUR models are divided, with the structural model indicating an increase, and the approximate system-theoretic model indicating a decrease. All six models forecast an increase in red king crab biomass in 1989. Figures 12 through 16 are graphical representations of the performance of the six models on the five series. Management authority estimates of biomass are denoted by crosses, and model forecasts and errors are indicated by continuous lines.

Despite the very limited number of annual biomass and catch observations, the multispecies models of walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab were remarkably successful. The MLE, CIDLS, MLE_{STTS}, and $CIDLS_{STTS}$ models reduced the percentage root-mean-squared error in the halibut equation to little over half the %RMSE obtained for the univariate and trivariate models (Table 6 (Area 4) and Table 14, respectively). The same four multispecies models also achieved a better fit in the king crab equation than was obtained by any of the earlier multispecies models. Overall fit in the yellowfin sole equation was not as good as in less complex models, (see Tables 2 and 14). The improved fit in the halibut and crab equations, accompanied by the relatively small deterioration in the fit of the yellowfin sole equation, speaks strongly in favor of a multispecies modeling approach.

Table 15. -- Biomass and landings of walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab (1000 t).

	Pollock	Halibut	Cod	Sole	Crab
Biomass					
1979	7800	3.35	792.3	1655	127.71
1980	9100	2.94	913.3	1606.5	101.74
1981	9900	2.73	840.1	2051.5	32.36
1982	10500	2.95	1013.9	2609.6	11.96
1983	9900	3.41	1126.4	3688.1	3.68
1984	9600	4.17	999.7	3136.8	7.33
1985	8400	5.21	957.6	2122.4	5.91
1986	6700	5.73	1134.1	1787.7	14.48
1987	7645	5.76	1142.4	2367.8	20.83
1988	6998	5.58	898.6	2741.5	25.36
Catch					
1979	923.4	0.62	39.4	99.0	49.01
1980	1016.4	0.32	51.6	87.4	59.07
1981	1029.0	0.54	62.5	97.3	15.27
1982	1013.9	0.19	66.6	95.7	1.36
1983	1041.4	0.90	93.2	108.4	0.0
1984	1180.6	0.97	133.2	159.5	1.90
1985	1237.5	1.23	145.4	227.2	1.90
1986	1235.6	1.45	140.8	208.6	5.18
1987	1234.9	2.03	144.4	181.4	5.59
1988	1316.5	2.08	192.7	223.2	3.36

Pollock (Wespestad and Traynor 1988). halibut (Bob Trumble. IPHC. Seattle WA. pers. commun.), cod (Thompson and Shimada 1988). yellowfin sole (Bakkala and Wespestad 1988). red king crab (Jerry Reeves, Alaska Fisheries Science Center, Seattle WA., pers. commun.).

Table 16. -- Multispecies population dynamics models of the Bering Sea: Coefficients.

Maximum Likelihood (MLE)

$$x_t = \begin{bmatrix} .80^* & -.56^* & 6.35^* & -.72 & 12.14^* \\ -.02 & 1.49^* & -2.04^* & .68^* & -4.99^* \\ .10^* & .10^* & .14 & -.11 & -1.27 \\ .07 & -.46^* & 5.03^* & -.38 & -7.04^* \\ -.003 & .01^* & -.04 & .01^* & 1.16^* \end{bmatrix} x_{t-1} - h_{t-1} + \varepsilon_t^{MLE}$$

Constrained Indirect Least Squares (CIDLS)

$$x_t = \begin{bmatrix} 1.04^* & & & & 16.92^* \\ & 1.45^* & -1.69^* & .55^* & -7.61^* \\ .12^* & .18^* & -.72^* & & \\ & -.39^* & 4.24^* & -.01 & \\ -.002^* & .01^* & -.10^* & .02^* & 1.30^* \end{bmatrix} x_{t-1} - h_{t-1} + \varepsilon_t^{CIDLS}$$

Seemingly Unrelated Regression (SUR)

$$x_t = \begin{bmatrix} 1.11^* & & & & \\ & 1.30^* & & & \\ & & 1.10^* & & \\ & & & 1.07^* & \\ & & & & 1.07 \end{bmatrix} x_{t-1} - h_{t-1} + \varepsilon_t^{SUR}$$

System-theoretic Time-series Model of MLE Errors ($STTS_{MLE}$)

$$\varepsilon_t^{MLE} = \begin{bmatrix} -0.07 & -0.04 & 0.03 & 0.01 & 0.002 \end{bmatrix}' z_{t|t-1} + e_t$$

$$z_{t+1|t} = -0.57 z_{t|t-1} + \begin{bmatrix} -13.18 & 103.03 & 144.02 & 34.89 & -550.42 \end{bmatrix} e_t$$

System-theoretic Time-series Model of CIDLS Errors ($STTS_{CIDLS}$)

$$\varepsilon_t^{CIDLS} = \begin{bmatrix} -0.33 & 0.08 & -0.07 & -0.17 & -0.01 \end{bmatrix}' z_{t|t-1} + e_t$$

$$z_{t+1|t} = 0.15 z_{t|t-1} + \begin{bmatrix} 22.98 & -192.55 & -244.23 & -59.62 & 912.19 \end{bmatrix} e_t$$

System-theoretic Time-series Model of SUR Errors ($STTS_{SUR}$)

$$\varepsilon_t^{SUR} = \begin{bmatrix} 0.34 & -0.22 & -0.03 & 0.43 & -0.01 \end{bmatrix}' z_{t|t-1} + e_t$$

$$z_{t+1|t} = 0.40 z_{t|t-1} + \begin{bmatrix} 0.21 & -1.00 & 3.04 & 0.15 & -24.19 \end{bmatrix} e_t$$

The vector x_t is ordered: walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab. Pacific halibut are measured in thousands of metric tons. All other species are measured in millions of metric tons. See Table 15 for the data and a description of their sources. Coefficients significant at the 95% level are denoted by an asterisk.

Table 17. -- Multispecies population dynamics models of the Bering Sea: Statistics.

	MLE	CIDLS	SUR	MLE_{STTS}	$CIDLS_{STTS}$	SUR_{STTS}
R^2						
Pollock	0.852	0.712	0.455	0.869	0.655	0.548
Halibut	0.992	0.987	0.837	0.993	0.990	0.867
Cod	0.544	0.181	-0.205	0.557	0.493	-0.086
Sole	0.848	0.717	0.094	0.847	0.750	0.467
King Crab	0.985	0.965	0.946	0.984	0.978	0.962
Percentage Root-Mean-Squared Error						
Pollock	5.6	7.8	10.8	5.3	7.2	9.8
Halibut	2.6	3.3	11.6	2.4	2.9	10.4
Cod	8.1	10.9	13.2	8.0	8.6	12.5
Sole	10.5	14.4	25.7	10.6	13.5	19.7
King Crab	14.3	21.9	27.2	15.0	17.3	23.0
Mean Absolute Deviation						
Pollock	0.405	0.487	0.793	0.381	0.485	0.713
Halibut	0.083	0.102	0.377	0.076	0.086	0.332
Cod	0.064	0.082	0.105	0.061	0.066	0.104
Sole	0.211	0.261	0.507	0.214	0.243	0.333
King Crab	0.004	0.006	0.008	0.004	0.005	0.007
Ljung-Box Q						
Pollock	12.43	3.09	2.79	12.01	7.63	3.21
Halibut	11.45	6.65	6.56	9.02	3.50	4.90
Cod	4.01	5.54	9.24	3.29	1.19	9.31
Sole	2.76	7.28	9.80	2.62	4.32	4.72
King Crab	6.28	10.00	1.64	6.62	4.63	2.78
Log Likelihood	99.9	93.7	16.7	104.0	98.6	22.6
Schwartz's Criterion	-102.0	-120.9	-13.8	-8.4	-29.0	76.2
Out-of-sample Forecasts (1989)						
Pollock	5.168	6.395	6.431	5.150	6.462	6.144
Halibut	5.965	5.811	5.165	5.949	5.813	5.354
Cod	0.847	1.029	0.796	0.860	1.036	0.839
Sole	0.982	1.395	2.701	0.985	1.396	2.360
King Crab	0.050	0.065	0.024	0.052	0.066	0.030

Halibut are reported in thousand t units, all other species are in million t. Survey estimates of 1989 biomass are not available! at the time of this writing. MLE is the unrestricted first-order linear model. CIDLS is the constrained indirect least-squares first-order linear model. SUR is the seemingly unrelated first-order linear model. MLE_{STTS} is the unrestricted first-order linear model with time series treatment of the residuals. $CIDLS_{STTS}$ is the constrained indirect least-squares first-order linear model with time series treatment of the residuals. SUR_{STTS} is the seemingly unrelated first-order linear model with time series treatment of the residuals.

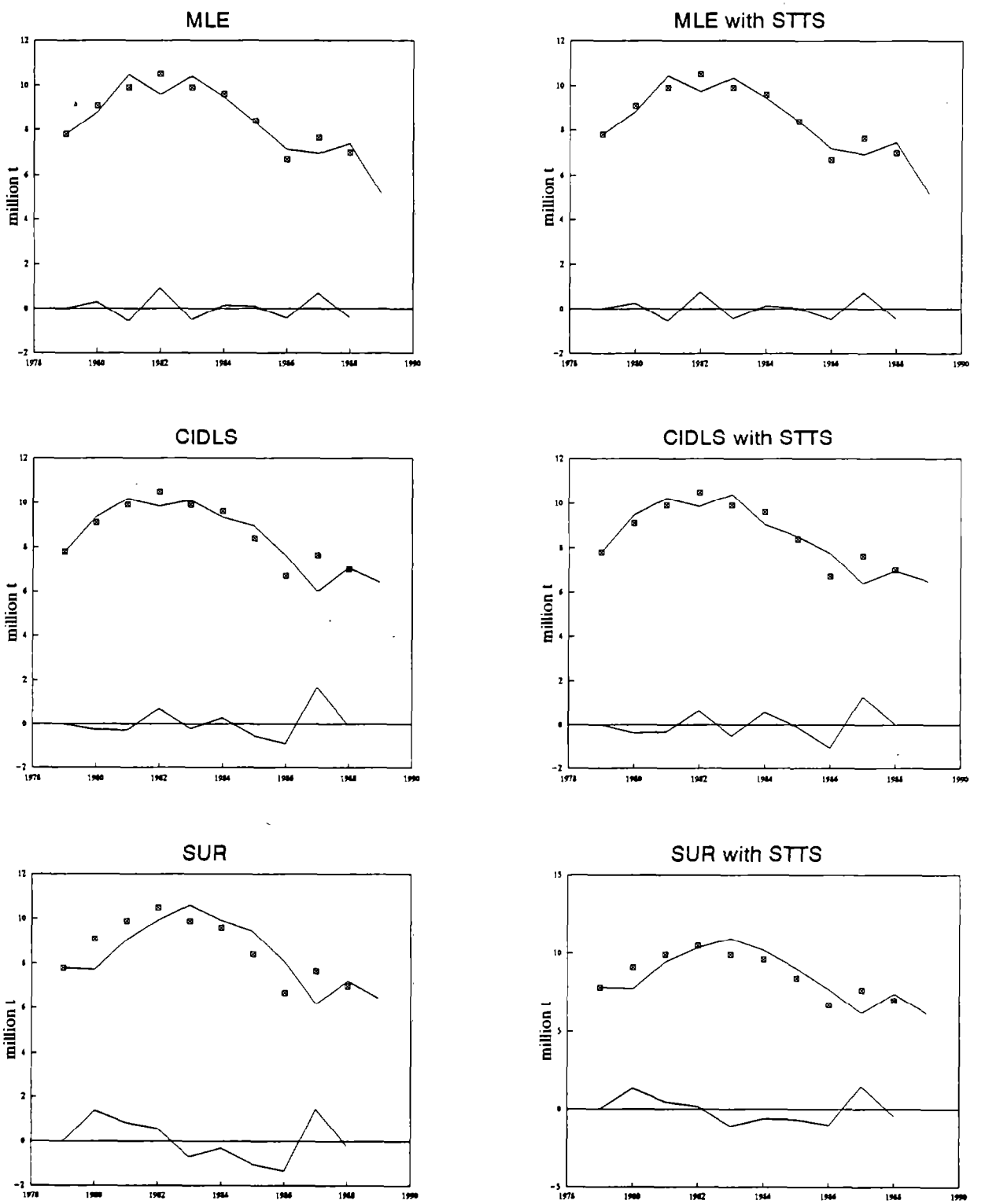

Figure 12.-- Six models of walleye pollock biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

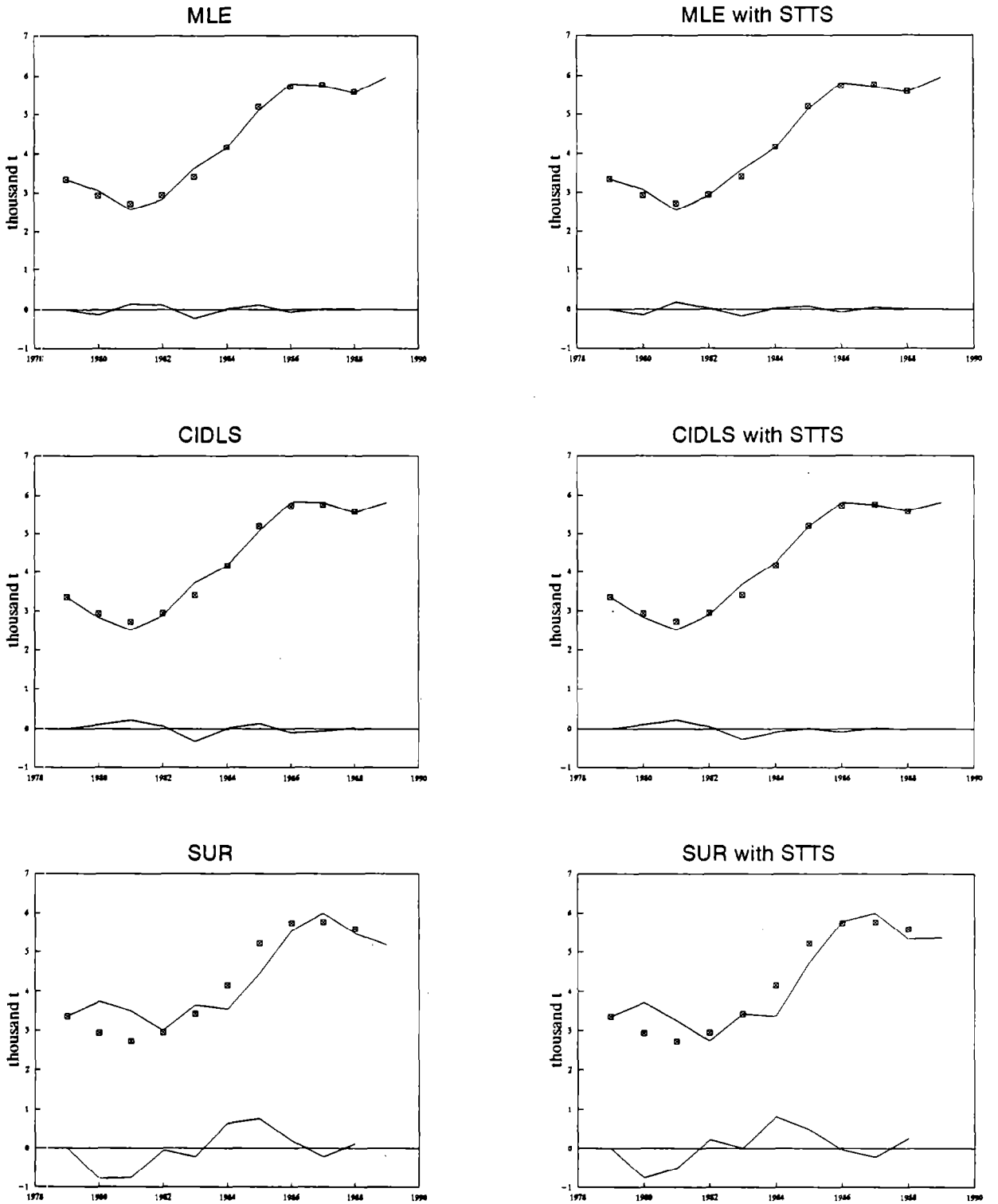

Figure 13. -- Six models of Pacific halibut biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

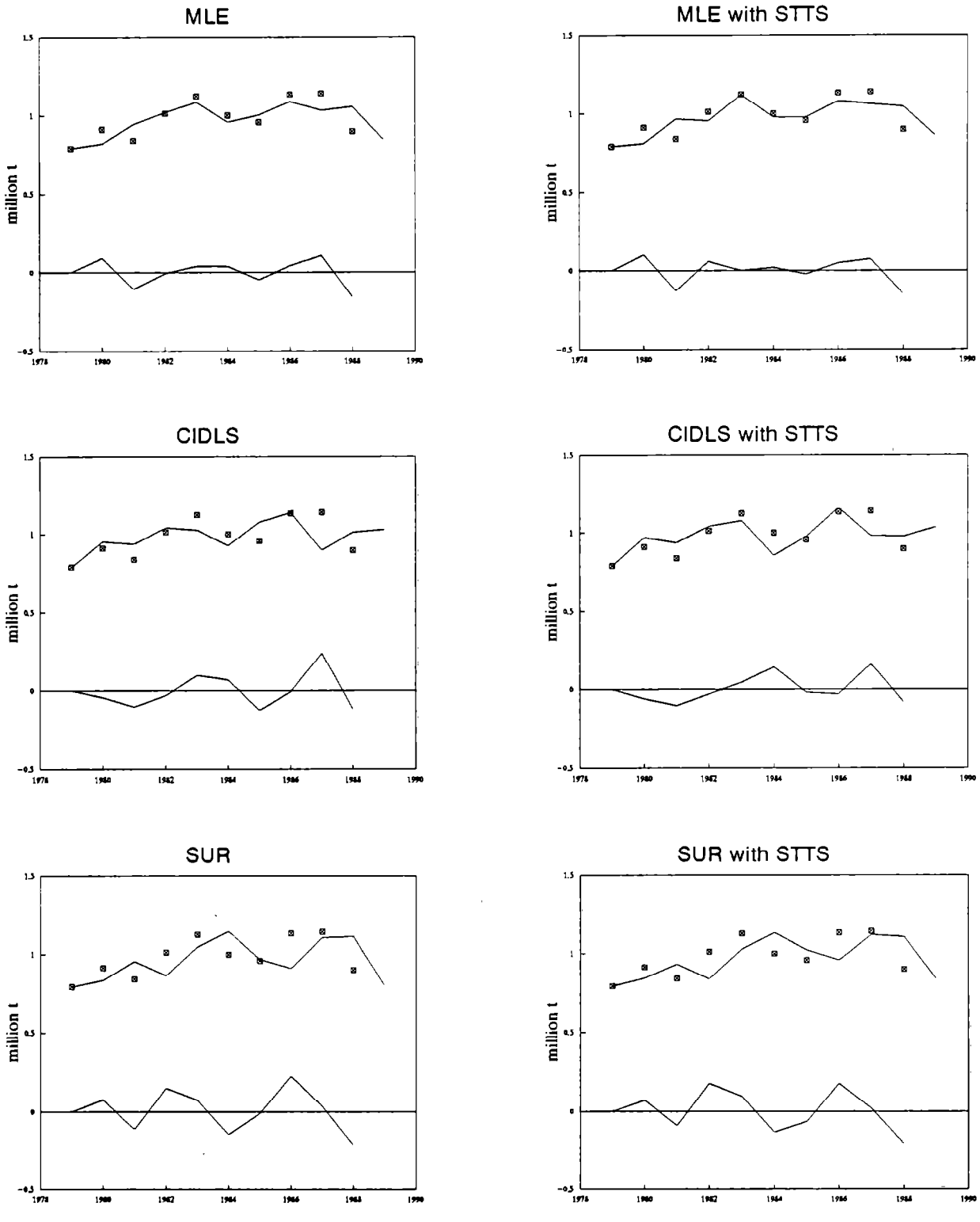

Figure 14. --Six models of Pacific cod biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

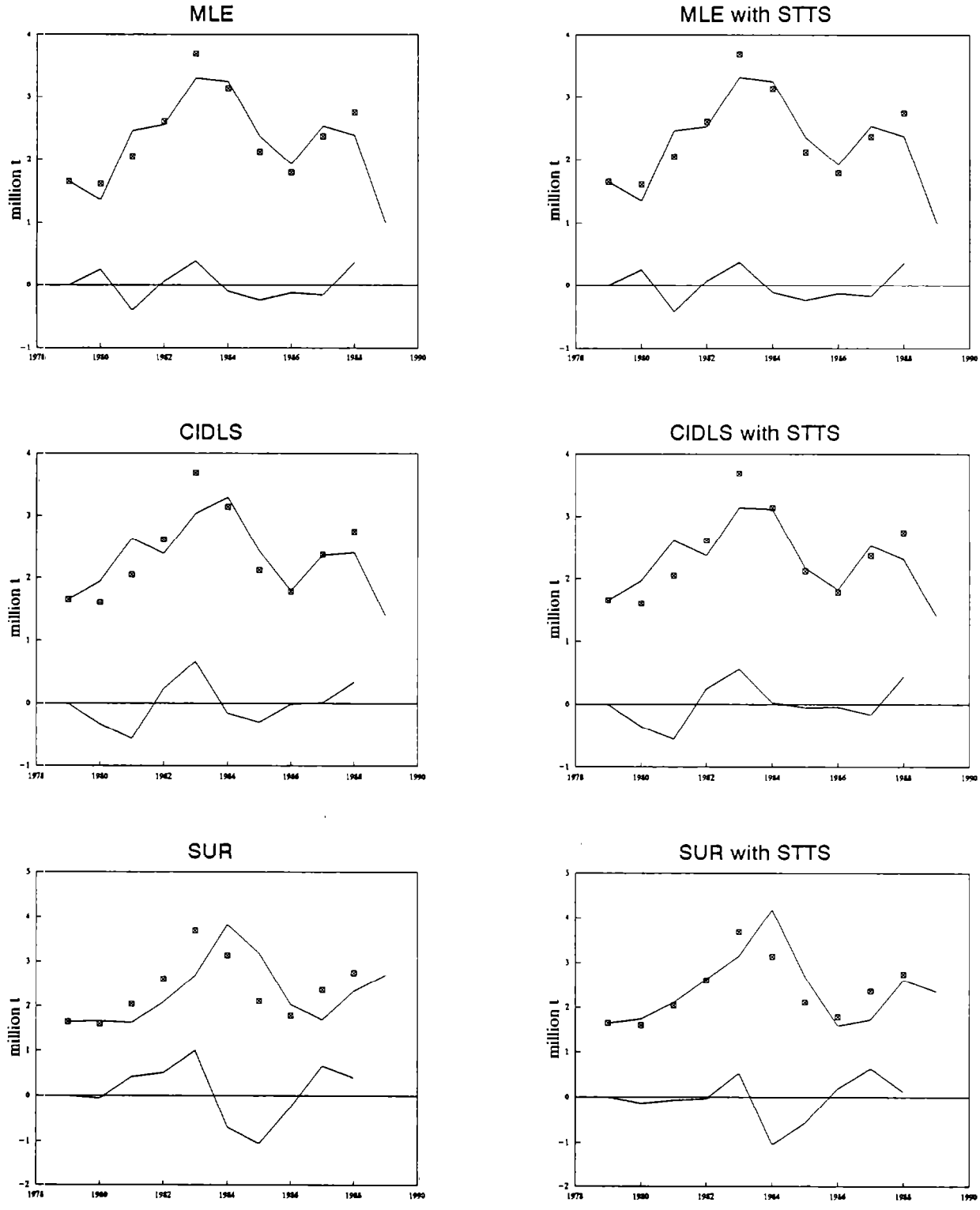

Figure 15. -- Six models of yellowfin sole biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

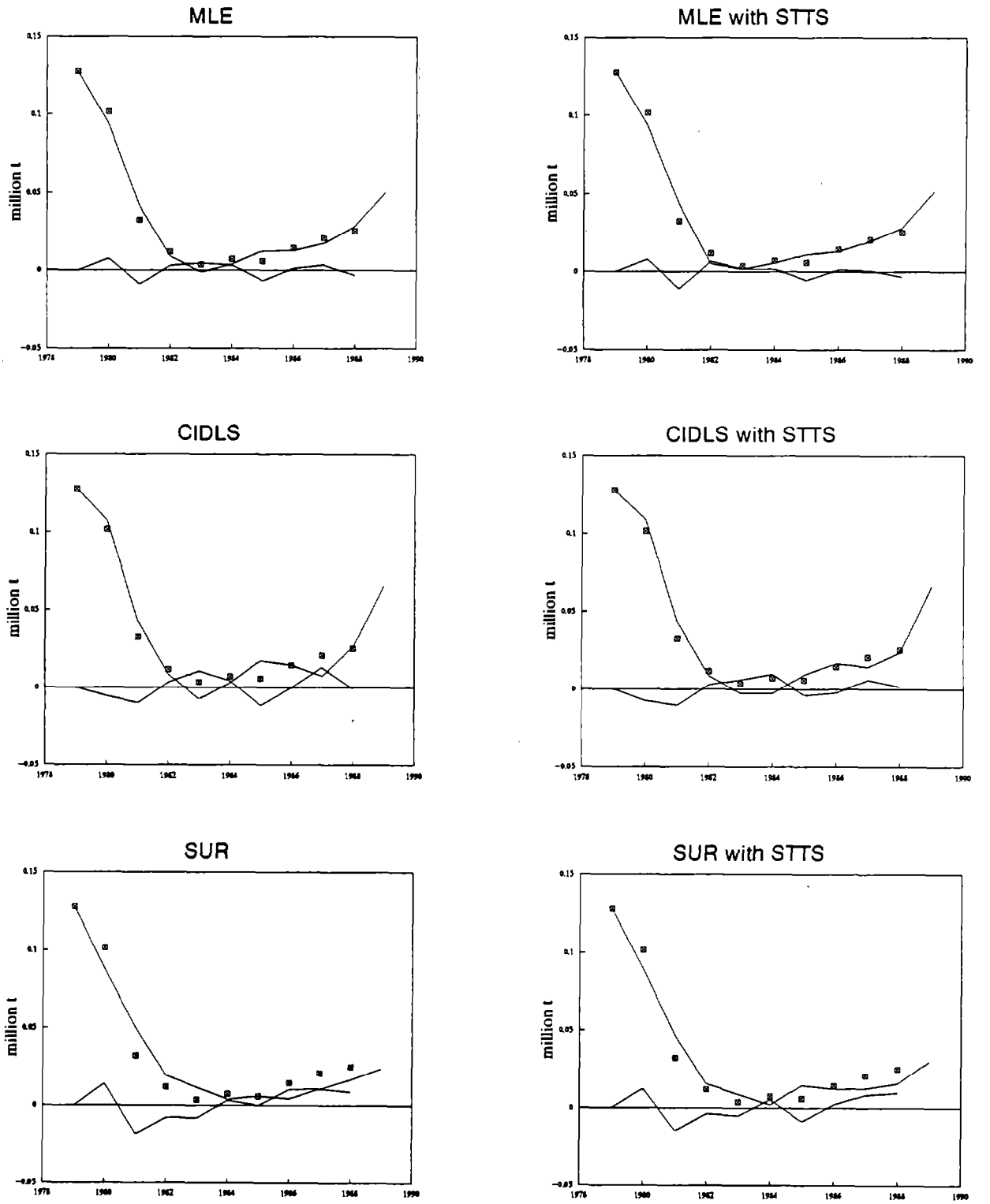

Figure 16. -- Six models of red king crab biomass. Observed biomass is denoted by crossed boxes. Estimates and residuals are represented by solid lines.

CONCLUSION

Model specification is the most difficult aspect of modeling dynamic systems. As a result, interesting dynamic systems are often characterized by a large number of mutually exclusive representations. A nonlinear encompassing model for population dynamics was developed in the first part of this document. By exploiting well-known mathematical results, it was demonstrated that the encompassing model could be approximated using dynamic polynomial expansions.

Approximation of the encompassing model involves the approximation of both nonlinear and dynamic relationships among the variables. It was demonstrated that the search for the maximum polynomial degree supported by the data is conditionally one-dimensional given a specification of model dynamics. For models of population dynamics, it is possible to determine the relevance of omitted higher degree polynomial terms through tests on the significance of the intercept. In addition to demonstrating the technique, the univariate yellowfin sole examples produce useful forecasts and suggest interesting policy conclusions. It was found that the most important dynamics of yellowfin sole populations are included in a linear representation. Quadratic models of yellowfin sole biomass explain less than 1% more of the observed variation than is explained by a simple linear model. Although the long run properties of linear and quadratic models are quite different, over the short time horizon relevant to fisheries management, the models are virtually indistinguishable.

The shortness of biological data-sets and the variety of possible interactions makes estimation of multispecies models difficult. Several common structural specifications can be motivated as a priori restrictions on dynamic polynomial expansions of the general encompassing model. Rather than asserting these restrictions, it is suggested that restrictions be selected using statistical techniques. A three-species system (Pacific halibut, yellowfin sole, and red king crab) was used to contrast the performance of restricted and unrestricted models. Careful selection of restrictions using econometric methods was shown to result in a model preferred over the unrestricted model on the basis of Schwartz's criterion.

It is possible to model dynamic relationships with time-series techniques as long as the processes generating the series are stationary. Because even nonlinear processes have linear state-space representations, a state-space time-series approach can be used to model population dynamics. A recently introduced multivariate time-series modeling algorithm which is particularly suited to data-based specification of model dynamics was reviewed and extended. Several time series models of halibut biomass were reported. These models indicate that halibut population dynamics are very similar across their entire range. However, the models suggest that harvest limits should still be set on as fine a spatial scale as possible because of local differences in initial conditions.

In the study of dynamic nonlinear systems, the researcher is often faced with a dilemma: on the one hand, the structural approach required for policy analysis often fails to yield useful forecasts, while on the other hand, although time-series methods may provide more accurate forecasts, they are uninformative about policy. The approximate structural/state-space methodology results in a model specification algorithm that is closely tied to the data while retaining the statistically supportable structural relationships required for policy analysis. The strength of the methodology is established with several examples. The first application concludes that the decline in coastwide Pacific halibut biomass during the 1960s and early 1970s, and its subsequent recovery, can be explained largely as the result of management policy. The second application revisited the three-species structural model. Models including a time-series treatment of the structural residuals resulted in substantial reductions in percentage root-mean-squared error, but were not preferred by Schwartz' criterion because of the increased number of coefficients and limited sample size. The third application models the important walleye pollock, Pacific halibut, Pacific cod, yellowfin sole, and red king crab fisheries of Bristol Bay and the eastern Bering Sea. The model demonstrates the importance of a multispecies approach and illustrates that the approximate structural/time-series methodology can even be applied to short time-series of data with some success.

CITATIONS

Aoki, M. 1987. State-space modeling of time-series. Springer-Verlag, New York, 314 p.

Aoki, M., and A. Havenner. 1989. A method for approximate representation of vector-valued time-series and its relation to two alternatives. J. Econometrics 42(2): 181-199.

Aoki, M., and A. Havenner. 1991. State-space modeling of multivariate time-series. Econometric Rev. (in press for 1991).

Armstrong, M. J., and P. A. Shelton. 1988. Bias in the estimation of stock-recruit function parameters caused by nonrandom environmental variability. Can. J. Fish. Aquat. Sci. 45:554-557.

Bakkala, R. G., and V. G. Wespestad. 1986. Yellowfin Sole. In R. G. Bakkala and L-L. Low, (editors), Condition of groundfish resources of the eastern Bering Sea and Aleutian Island Regions in 1985, p. 49-62. U.S. Dep. Commer., NOAA Tech. Memo. NMFS F/NWC-104.

Bakkala, R. G. and V. G. Wespestad. 1988. Yellowfin Sole. In R. G. Bakkala and L-L. Low (editors), Condition of groundfish resources of the eastern Bering Sea and Aleutian Islands Region in 1987, p. 51-63. U.S. Dep. Commer., NOAA Tech. Memo. NMFS F/NWC-139.

Berck, P., and G. Johns. 1985. Policy consequences of better stock estimates in Pacific halibut fisheries. Proceedings of Bus. Econ. Stat. Section of the American Statistical Association, p. 139-145.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Minist. Agric. Fish. Food (GB) Ser. II 19:533 p.

Cohen, Y., and J. N. Stone. -1987. Multivariate time-series analysis of the Canadian fisheries system in Lake Superior. Can. J. Fish. Aquat. Sci. 44 (Supplement 2):171-181.

Criddle, K. R., and A. Havenner. 1989. Forecasting halibut biomass using system-theoretic time-series methods. Am. J. Ag. Econ. 71(2):422-431.

Criddle, K. R. and A. Havenner. 1991. An encompassing approach to modeling fishery dynamics: Modeling dynamic nonlinear systems. Nat. Res. Model. (in press for 1991).

Fogarty, M. J. 1988. Time-series models of the Maine lobster fishery: The effects of temperature. Can. J. Fish. Aquat. Sci. 45: 1145-1153.

Havenner, A., and M. Aoki. 1988b. Deterministic and stochastic trends in state-space models of nonstationary time-series. Univ. California, Davis, Dep. Agricultural Economics, Working Paper 88-17. Department of Agricultural Economics, University of California Davis, Davis, CA.

Havenner, A., and M. Aoki. 1988a. An instrumental variable interpretation of linear system theory estimation. J. Econ. Dynamics Control 12:49-54.

Havenner, A., and M. Aoki. 1988c. Econometrics and linear system theory in multivariate time-series analysis. Univ. California, Davis, Dep. Agricultural Economics, Working Paper 88-4. Department of Agricultural Economics, University of California Davis, Davis, CA.

Havenner, A., and K. R. Criddle. 1989. System-theoretic time-series: An application to inventories and prices of range cattle in California. Computers Mathemat. Applica. 17:1177-1188.

International Pacific Halibut Commission. 1987. The Pacific halibut: Biology, fishery and management. Tech. Rep. 22, International Pacific Halibut Commission, P.O.Box 95009, Seattle, WA 98145.

International Pacific Halibut Commission. 1988. Annual Report 1987. International Pacific Halibut Commission, P.O.Box 95009, Seattle, WA 98145.

Kailath, T. 1980. Linear systems. Prentice-Hall, Englewood Cliffs, New Jersey.

Lehmann, E. L. 1959. Testing statistical hypotheses, Wiley, New York.

Livingston, P. A., and B. J. Goiney, Jr. 1983. Food habits literature of North Pacific marine fishes: A review and selected bibliography. U.S. Dep. Commer., NOAA, Tech. Memo. NMFS F/NWC-54, 81 p.

Mendelssohn, R. 1980. Using Box-Jenkins models to forecast fishery dynamics: Identification, estimation, and checking. Fish. Bull., U.S. 78:887-896.

Noakes, D., D. W. Welch and M. Stocker, 1987. A time-series approach to stock recruitment analysis: Transfer function noise modeling. Nat. Res. Model. 2:213-233.

Paganol, M., and M. J. Hartley. 1981. On fitting distributed lag models subject to polynomial restrictions. J. Econometrics 16:171-198.

Quinn II, T. J., R. B. Deriso, and S. H. Hoag. 1985. Methods of population assessment of Pacific halibut. International Pacific Halibut Commission, Scientific Report 72. IPHC, P.O. Box 95009, Seattle, WA 98145.

Ricker, W. E. 1954. Stock and recruitment. J. Fish. Res. Board of Can. 11:559-623.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Can. Bull. Fish. Aquat. Sci. 191.

Saila, S. B., M. Wigbout and R. J. Lermit. 1980. Comparisons of some time-series models for the analysis of fisheries data. J. Con. Int. Explor. Mer 39:44-52.

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull. Inter-Am. Trop. Tuna Comm. 1(2):27-56.

Stevens, B. G., and R. A. Macintosh. 1986. Report to industry on the 1986 eastern Bering Sea crab survey. NWAFC Processed Rep. 86-20, 52 p,, Northwest and Alaska Fish. Cen., Natl. Mar. Fish. Ser., NOAA, 7600 Sand Point Way NE., BIN C15700, Seattle, WA 98115.

Stevens, B. G., R. A. MacIntosh, and K. L. Stahl-Johnson. 1988. Report to industry on the 1988 eastern Bering Sea crab survey. NWAFC Processed Rep. 88-23,41 p. Northwest and Alaska Fish. Cent., Natl. Mar. Fish. Ser., NOAA, 7600 Sand Point Way NE., BIN C15700, Seattle, WA 98115.

Thompson, G. G., and A. M. Shimada. 1988. Pacific cod. In R. G. Bakkala (editor), Condition of groundfish resources of the eastern Bering Sea and Aleutian Islands Region in 1987, p. 33-49, U.S. Dep Commer., NOAA Tech. Memo. F/NWC-139.

Trivedi, P. K., and A. R. Pagan. 1979. Polynomial distributed lags: A unified approach. Econ. Stud. Quart. 30:37-49.

Uhler, R. S. 1979. Least squares regression estimates of the Schaefer production model: Some Monte Carlo results. Can. J. Fish. Aquat. Sci. 37: 1284-1294.

Volterra, V. 1928. Variations and fluctuations of the number of individuals in animal species living together. J. Con. Int. Explor. Mer 3:3-51.

Walters, C. J. 1985. Bias in the estimation of functional relationships from time-series data. Can. J. Fish. Aquat. Sci. 42:147-149.

Walters, C. J. 1987. Nonstationarity of production relationships in exploited populations. Can. J. Fish. Aquat. Sci. 44 (Supplement 2):156-165.

Wegge, L. L. 1978. Constrained indirect least squares estimators. Econometrica 46(2):435-449.

Wespestad, V. G., and J. J. Traynor. 1988. Walleye Pollock. In R. G. Bakkala (editor), Condition of groundfish resources of the eastern Bering Sea and Aleutian Islands Region in 1987, p. 11-32. U. S. Dep. Commer., NOAA Tech. Memo. NMFS F/NWC-139.

Zellner, A., and F. Palm. 1974. Time-series analysis and simultaneous equation models. J. Econometrics 2: 17-54.

ACKNOWLEDGMENTS

This research was supported in part by the Resource, Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA. Opinions and conclusions are the responsibility of the author.

I wish to express appreciation to Art Havenner, Dick Howitt, and Jim Wilen for their support, advice, and assistance. The painful task of preparing this manuscript was cheerfully undertaken by Mary Parsons.

APPENDIX

Examples of Nested Structural Models

The model presented as Equation (5) can be shown to encompass many structural models. This section presents first- and second-degree expansions of four structural specifications: the logistic surplus yield (Schaefer 1954), a delay-difference model with Ricker recruitment, a delay-difference model with Beverton-Hole recruitment, and a generalized Lotka-Volterra model.

Single-Species Models

Schaefer--Schaefer (1954) is generally credited with popularization of the Pearl-Verhulst logistic surplus yield model in fisheries. A discrete-time representation of the logistic surplus yield model is

$$x_t = [1+g]x_{t-1} - \frac{g}{k}x_{t-1}^2 - h_{t-1},$$

where g and k are the intrinsic growth rate and carrying capacity, respectively. The Jacobian matrix is a row vector of the first partial derivatives of the model with respect to x_t, x_{t-1}, and h_{t-1}, i.e.,

$$\mathcal{J} = [1, 2\frac{g}{k}x_{t-1} - [1+g], 1] .$$

Applying Taylor's expansion, a linear first-order approximation of the model is

$$\mathcal{f}(x) \cong x_0 - [1+g]x_0 + \frac{g}{k}x_0^2 + h_0 + x_t - x_0 + [x_{t-1}-x_0][2\frac{g}{k}x_0-1-g] + h_{t-1} - h_0,$$

which on cancellation reduces to

$$x_t \cong [1+g]x_{t-1} - 2\frac{g}{k}x_0x_{t-1} + \frac{g}{k}x_0^2 - h_{t-1} .$$

Because the expansion point is predetermined, the first-order approximate logistic model is a linear equation in x_{t-1}

$$x_t \cong \beta_0 + \beta_1 x_{t-1} - h_{t-1}.$$

Because the logistic model is quadratic in x_{t-1}, a second-degree Taylor-series expansion will also be exact. The quadratic Taylor-series expansion is

$$f(x) = x_0 - [1+g]x_0 + \frac{g}{k}x_0^2 + h_0 + \begin{bmatrix} x_t-x_0 & & \\ & x_{t-1}-x_0 & \\ & & h_{t-1}-h_0 \end{bmatrix} \mathcal{J}$$

$$+ \frac{1}{2}\begin{bmatrix} x_t-x_0 & & \\ & x_{t-1}-x_0 & \\ & & h_{t-1}-h_0 \end{bmatrix} \mathcal{H} \begin{bmatrix} x_t-x_0 \\ x_{t-1}-x_0 \\ h_{t-1}-h_0 \end{bmatrix},$$

where $\mathcal{H}$ is the Hessian or matrix of second partial derivatives:

$$\mathcal{H} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 2\frac{g}{k} & 0 \\ 0 & 0 & 0 \end{vmatrix}.$$

After introducing the coefficients of the Jacobian and Hessian, the first-order quadratic expansion of the Schaefer model can be rewritten as

$$f(x) = x_0 - [1+g]x_0 + \frac{g}{k}x_0^2 + h_0 + [x_t-x_0] + [x_{t-1}-x_0][2\frac{g}{k}x_0 - (1+g)] + [h_{t-1}-h_0]$$

$$+ \frac{1}{2}[x_{t-1}-x_0]^2 2\frac{g}{k},$$

which on cancellation reduces to

$$x_t = [1+g]x_{t-1} - \frac{g}{k}x_{t-1}^2 - h_{t-1},$$

the discrete Schaefer model.

In general, increasingly high-degree polynomial expansions become arbitrarily close approximations to the true nonlinear process, and are exact when the degree of the approximation is equal! to the degree of the system.

Ricker--Ricker (1954) introduces a model of population dynamics that includes a function of lagged abundance to represent recruitment. Following Ricker (1975, p. 283) the stock-recruitment function can be modeled as

$$f(x_{t-\ell}) = x_{t-\ell}\{\exp[g_\ell[1-\frac{x_{t-\ell}}{k_\ell}]]\}$$

$$\cong \{1 + g_\ell[1-\frac{x_{t-\ell}}{k_\ell}]\}x_{t-\ell}.$$

When current net-growth functions are also modeled as logistic, the Ricker delay-difference model is an extension of the Schaefer model:

$$x_t \cong [1+g_1]x_{t-1} - \frac{g_1}{k_1}x_{t-1}^2 + [1+g_\ell]x_{t-\ell} - \frac{g_\ell}{k_\ell}x_{t-\ell}^2 - h_{t-1},$$

and a member of a more general family of delay-difference models

$$x_t \cong \sum_{i=1}^{\ell} [1+g_i]x_{t-i} - \frac{g_i}{k_i}x_{t-i}^2 - h_{t-1} ,$$

where g_i, and k_i are the intrinsic growth rates and carrying capacities at lag t-i. A first-degree Taylor-series approximation of the second order delay-difference model is

$$x_t \cong \sum_{i=1}^{\ell} [1+g_i]x_{t-i} - \frac{2g_i}{k_i}x_0x_{t-i} + \frac{g_i}{k_i}x_0^2 - h_{t-1} ,$$

or,

$$x_t \cong \beta_0 + \sum_{i=1}^{\ell}\beta_i x_{t-i} - h_{t-1} ,$$

a linear equation in x_{t-i}. First- and second-degree exponential Ricker models are similar, but do not have exact polynomial expansions because derivatives of exponential functions do not vanish.

Beverton-Holt--The population dynamics model introduced in Beverton and Holt (1957) is similar to the Ricker model described above, but has a stock-recruitment relationship that Ricker (1975, p. 291) represents as

$$f_\ell(x_{t-\ell}) = \frac{ax_{t-\ell}}{b+x_{t-\ell}} .$$

A Beverton-Holt delay-difference model with logistic growth is

$$x_t = [1+g]x_{t-1} - \frac{g}{k}x_{t-1}^2 + \frac{ax_{t-\ell}}{b+x_{t-\ell}} - h_{t-1},$$

with a first-degree Taylor-series expansion,

$$x_t \equiv \frac{g}{k}x_0 + \frac{ax_0}{[b+x_0]^2} + [1+g]\,x_{t-1} - \frac{2g}{k}\,x_0 x_{t-1} + \frac{ab}{[b+x_0]^2}\,x_{t-\ell} - h_{t-1},$$

a linear function of current and lagged abundance

$$x_t \equiv \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-\ell} - h_{t-1}\,.$$

The quadratic Taylor-series approximation of the first order Beverton-Holt delay-difference model is

$$x_t \equiv [1+g]x_{t-1} - \frac{g}{k}\,x_{t-1}^2 + \frac{ab[b+3x_0]}{[b+x_0]^3}\,x_{t-\ell} - \frac{ab}{[b+x_0]^3}\,x_{t-\ell}^2 + \frac{ax_0^3}{[b+x_0]^3} - h_{t-1}$$

or,

$$x_t \equiv \beta_0 + \beta_1 x_{t-1} - \beta_2 x_{t-1}^2 + \beta_3 x_{t-\ell} - \beta_4 x_{t-\ell}^2 - h_{t-1}\,.$$

Although second-degree expansions of Beverton-Holt models are not exact, omitted higher-degree terms are unimportant because of the rapidity with which the derivatives of the recruitment function attenuate."

Multispecies Models

Lotka-Volterra--Matrix models of population dynamics were developed independently by Lotka and Volterra. Following Volterra (1928), the multispecies logistic surplus yield model can be represented by

$$x_t = [I+G]x_{t-1} - D(x_{t-1})Kx_{t-1} + h_{t-1}\,.$$

Linear and quadratic net growth, predation, competition, and recruitment are accounted for in the s x s matrices G and K. The diagonal matrix $D(x_{t-1})$ is composed of the elements of x_{t-1}. The Jacobian matrix is

$$\mathcal{J} = [I,\ 2D_{x(t-1)}K - (I+G),\ I]$$

A first-degree Taylor-series expansion of the Lotka-Volterra model is

$$\mathcal{J}(x) = x_0 - [I+G]x_0 + D(x_0)K]x_0 + h_0 + [x_t - x_0] + [2D(x_0)K - [I+G]][x_{t-1} - x_0] + [h_{t-1} - h_0]$$

[11] The derivatives of the Beverton-Holt stock-recruitment function for values of a = 15.84 and b = 2.54 reported for North Sea plaice in Ricker [1975 p-293] evaluated at the mean stock level equal to 26 are: -0.0495, 0.00347, -0.000364. and 0.0000510 for the first through fourth derivatives, respectively.

which simplifies to

$$x_t \cong D(x_0)Kx_0 + [I+G]x_{t-1} - 2D(x_0)Kx_{t-1} + h_{t-1} ,$$

or,

$$x_t \cong \beta_0 + \beta_1 x_{t-1} - h_{t-1} ,$$

where B_0 is an s-element vector, and B_1 is an s x s matrix of coefficients.